



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 197086

TO: Ginny Portner
Location: REM/3B02/3C18
Art Unit: 1645
Monday, August 07, 2006
Case Serial Number: 09/732091

From: Vira David
Location: Biotech-Chem Library
REM-1A41
Phone: (571)272-1972

Virajita.David@uspto.gov

Search Notes

Examiner Portner,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Vira David
Intern
STIC Biotech/Chem Library
(571)272-1972

STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact **the searcher or contact:**

Mary Hale, Information Branch Supervisor
571-272-2507 Remsen E01 D86

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.

STIC-Biotech/ChemLib

197086

From: Portner, Ginny
Sent: Monday, July 31, 2006 1:04 PM
To: STIC-Biotech/ChemLib
Subject: 09/732,091

please interference search SEQ ID No 3 and 4. thanks!

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

7/31/2006

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2006, 09:12:51 ; Search time 202 Seconds
(without alignments)
7030.552 Million cell updates/sec

Title: US-09-732-091-3

Perfect score: 759

Sequence: 1 atggcatacaaatatgatag.....cgttgcaatagaatccatt 759

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /EMC Celleria_SIDS3/ptodata/2/ina/1 COMB.seq:*
- 2: /EMC Celleria_SIDS3/ptodata/2/ina/5 COMB.seq:*
- 3: /EMC Celleria_SIDS3/ptodata/2/ina/6A COMB.seq:*
- 4: /EMC Celleria_SIDS3/ptodata/2/ina/6B COMB.seq:*
- 5: /EMC Celleria_SIDS3/ptodata/2/ina/7 COMB.seq:*
- 6: /EMC Celleria_SIDS3/ptodata/2/ina/H COMB.seq:*
- 7: /EMC Celleria_SIDS3/ptodata/2/ina/PCTUS COMB.seq:*
- 8: /EMC Celleria_SIDS3/ptodata/2/ina/PP COMB.seq:*
- 9: /EMC Celleria_SIDS3/ptodata/2/ina/RE COMB.seq:*
- 10: /EMC Celleria_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	68.2	9.0	7218	2	US-08-232-463-14
2	58.8	7.7	966	3	Sequence 14, Appl
3	43.4	5.7	822	3	Sequence 13269, A
4	41.4	5.5	183	3	Sequence 13270, A
5	41.4	5.5	3666	3	Sequence 1310, Ap
6	41.2	5.4	612	3	Sequence 3543, Ap
7	40.8	5.4	19124	2	Sequence 1357, Ap
8	40.2	5.3	5852	2	Sequence 13, Appl
9	40.2	5.3	640681	3	Sequence 2, Appl
10	40.2	5.3	1664976	3	Sequence 1, Appl
11	40.2	5.3	1305	3	Sequence 1, Appl
12	40	5.2	795	3	Sequence 65, Appl
13	39.8	5.2	525	3	Sequence 878, App
14	39.2	5.2	601	3	Sequence 14352, A
15	39.2	5.2	601	3	Sequence 17780, A
16	39.2	5.2	601	3	Sequence 88247, A
17	39.2	5.2	612	3	Sequence 75, Appl
18	39.2	5.2	612	3	Sequence 75, Appl
19	39.2	5.2	612	3	Sequence 75, Appl
20	39.2	5.2	612	3	Sequence 75, Appl
21	39.2	5.2	612	3	Sequence 75, Appl
22	39.2	5.2	612	3	Sequence 75, Appl
23	39.2	5.2	612	3	Sequence 75, Appl

24 39.2 5.2 612 3 US-09-834-759-75 Sequence 75, Appl
 25 39.2 5.2 612 3 US-09-590-751A-75 Sequence 75, Appl
 26 39.2 5.2 612 3 US-09-551-621-75 Sequence 75, Appl
 27 39.2 5.2 612 3 US-09-551-621A-75 Sequence 75, Appl
 28 39.2 5.2 612 3 US-10-076-622-75 Sequence 75, Appl
 29 39.2 5.2 612 4 US-10-124-805-75 Sequence 24, Appl
 30 39.2 5.2 1141 3 US-09-806-708B-22 Sequence 24, Appl
 31 39.2 5.2 5418 3 US-09-949-016-12477 Sequence 2477, Ap
 32 39.2 5.2 40315 3 US-09-949-016-11753 Sequence 11753, A
 33 39.2 5.2 40649 3 US-09-949-016-14219 Sequence 14219, A
 34 38.8 5.1 3332 3 US-09-710-279-3928 Sequence 3928, Ap
 35 38.8 5.1 3532 3 US-09-710-279-3752 Sequence 3752, Ap
 36 38.4 5.1 51621 3 US-09-949-016-12848 Sequence 12848, A
 37 38.4 5.1 51621 3 US-09-949-016-16503 Sequence 16503, A
 38 38.4 5.1 263693 3 US-09-949-016-12386 Sequence 12386, A
 39 38.4 5.1 263694 3 US-09-949-016-16915 Sequence 16915, A
 40 38.2 5.0 1956 3 US-08-559-896B-1 Sequence 1, Appli
 41 38.2 5.0 1956 3 US-09-351-794A-1 Sequence 1, Appli
 42 38.2 5.0 2823 3 US-10-104-047-169 Sequence 169, App
 43 38.2 5.0 3958 4 US-10-094-749-944 Sequence 944, App
 44 38 5.0 1039 3 US-09-902-540-1280 Sequence 1280, Ap
 45 38 5.0 6124 3 US-08-213-419B-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
 US-08-232-463-14/c
 ; Sequence 14, Application US/08232463
 ; Patent No. 5670367
 ; GENERAL INFORMATION:
 ; APPLICANT: DORNER, F.
 ; APPLICANT: SCHEIFLINGER, F.
 ; APPLICANT: FALKNER, F. G.
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,463
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/935,313
 ; FILING DATE:
 ; APPLICATION NUMBER: EP 91 114 300.6
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7218 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:

Db 65 ATTGTAGAAACAACTTATGAACAAACGCAACGCAACAAAGAGAAACATTTCTCTCAAGAAGAG 124
Qy 407 TATCCATAAAACAAACGACGACAAATTTAAACACAGACAA 441
Db 125 TATCTCAATGATTAAAGATCGTTAGCTAGAA 159

RESULT 5

US-09-710-279-3543
; Sequence 3543, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3543
; LENGTH: 3666
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3543

Query Match 5.5%; Score 41.4; DB 3; Length 3666;
Best Local Similarity 54.2%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 287 AGGTCAATTACAAACAGAAACTGAAACGACTTTAAATTGAACAAACATGCTTTCTAAAA 346
Db 2728 AGGTATATCTCATGAGAACAAATCAAAAGTAAATTTACTGAAGAACGAAACACAATGATA 2787
Qy 347 TCTTAGAAGAGTTTGAAGAAATGGATGATGAAGAGTGAAGAAATGTCGATGAAT 406
Db 2788 ATTGTAGAAACAACTTATGAACAAACGCAACAAAGAGAGAAACATTTCTCAAGAAGAG 2847
Qy 407 TATCCATAAAACAAACGACAAATTTAAACACAGCA 441
Db 2848 TATCTCAATGATTAAAGATCGTTAGCTAGAA 2882

RESULT 6

US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 5.4%; Score 41.2; DB 3; Length 612;
Best Local Similarity 44.6%; Pred. No. 0.15;
Matches 157; Conservative 0; Mismatches 195; Indels 0; Gaps 0;
Qy 90 TAAAGACGGCGGAAAGACACAAATGAAGAACTGACCAGCTCCATAGATAACAAAAGCA 149
Db 151 TAAAAAATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATA 210
Qy 150 TGGCGATGATTACGCTAAATATACGACAGAAAGAACTGCTGAAGAGTTGCAATATCTATGGAG 209
Db 211 AAAAAAATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATA 270
Qy 210 CAATAGTTTGGCGAGTTTCATTAAAGCGGAAGAGTCTTATACAAAGAGATTTATGCGA 269
Db 271 TAAAAAATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATA 330
Qy 270 TGTGTCGATTAATTAAGTCAATTTACACAGAAACTGAAACGACTTTAAATTGACA 329
Db 331 AAAAAAATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATA 390
Qy 330 AAACATGCTTTCTAAATCTTAGAAGAGAGTCTTGGAGAGAAATGATGATGAAGAGTCAA 389
Db 391 TAAAAAATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATA 450
Qy 390 AGAATGTGCGATGATTTATCCATAAAAAACACGACAAATTTAAACAGACAA 441
Db 451 TTAATAAATAAATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATA 502

RESULT 7

US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5933827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Welles, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

Query Match
Best Local Similarity 5.4%; Score 40.8; DB 2; Length 19124;
Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 278 ATAAATTAAGGTCAATTACACAGAGAACTGAAACGACTTTTAATTCACACAAACATGC 337
Db 15575 AAAAAATGAAAGAGATTATCAAAAAAAATTAATAAAAAATTTTATATAAAAAAAATGA 15634

QY 338 TTTCTAAATCTTAGAAGAGAGTTTGAAGAAATGGATGATGAAGAAGTGAAGAGAAATGT 397
Db 15635 TTATAAAAAATTAATAAAAAACAAAGAGAGAAAAACATTAATAAAAAAAATATAT 15694

QY 398 GCGATGAATTTATCCATATAAAAAACAGGACCAATTTTAAACAGACAA 441
Db 15695 ATCATAAAAACAAAAAAGAAAAAATATATTAATAATAAA 15738

RESULT 8
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; US-07-867-106-2

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-790-988-1/c

Query Match
Best Local Similarity 5.3%; Score 40.2; DB 3; Length 640681;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 193 TTGCAATCTACTATGGGAGCAATAGTTTTCGAGTTTCATTAAAGCCGAGAGTCTTATAC 252
Db 623982 TTATAATTTTATTAATAAAAAATATTTTTTATTTTAAAGACACAAAAAATATCTTTT 623923

QY 253 AAAGAGATTTTATGCGATGTCGCGATAAATTAAGGTCAATTAACACAGAAAACTGAA 312
Db 623922 TTGATGTTTATATATATTTTTTTTGGATTTAATATATATATTTTTTACAAAAATTAAT 623863

QY 313 ACGACTTTTAATGAACAAAAACATGCTTTCTTAAATCTTAGAAAGAGAGTTTGAAGAAATG 372
Db 623862 CTTTATTTGATTTGAATAGACATGCGTTTCAATATCTCCCAAAAAATTTTATTACTAC 623803

; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOHU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-790-988-1/c

Query Match
Best Local Similarity 5.3%; Score 40.2; DB 3; Length 640681;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 193 TTGCAATCTACTATGGGAGCAATAGTTTTCGAGTTTCATTAAAGCCGAGAGTCTTATAC 252
Db 623982 TTATAATTTTATTAATAAAAAATATTTTTTATTTTAAAGACACAAAAAATATCTTTT 623923

QY 253 AAAGAGATTTTATGCGATGTCGCGATAAATTAAGGTCAATTAACACAGAAAACTGAA 312
Db 623922 TTGATGTTTATATATATTTTTTTTGGATTTAATATATATATTTTTTACAAAAATTAAT 623863

QY 313 ACGACTTTTAATGAACAAAAACATGCTTTCTTAAATCTTAGAAAGAGAGTTTGAAGAAATG 372
Db 623862 CTTTATTTGATTTGAATAGACATGCGTTTCAATATCTCCCAAAAAATTTTATTACTAC 623803
```

Qy 373 GATGATGAAGAGTGAAGAATGTGCGATGAA 405
|||||
Db 623802 GATTACAATCCATTGATTATTTTATAATCAA 623770

RESULT 10
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g

```
; NAME/KEY: misc feature
; LOCATION: (111981)..(111981)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
;
; US-08-916-421B-1
;
; Query Match 5.3%; Score 40.2; DB 3; Length 1664976;
; Best Local Similarity 48.5%; Pred. No. 5.4;
; Matches 111; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
;
; Qy 103 AARAGACATGAAACTGACAGCTCCATAGAAAGGCGATGATTC 162
; Db 149587 AATAAATCAATGAAACAAAGCCAGAGTTGTTGATAATGATCACATGGGAGTTAAGAT 149528
;
; Qy 163 GCTAAATACGAGAAAGTCCGCTGAAGAGTTGCAATACTATGGAGCAATAGTTTTCG 222
; Db 149527 GACAAATCCTCAATAGAGGAGATATATTTCCAAATATCTGATAGAGTTATTC 149468
;
; Qy 223 AGTTTCATTAAGCGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGCGATAAA 282
; Db 149467 AGCAAGATTAGGACTGAAAGTTGAGCTTTTAAATGGAAGTATAAGTATCAGCTAATTAA 149408
;
; Qy 283 TTAAGTCAATTACACAGAAACGAAACGACTTTAATTGAACAAA 331
; Db 149407 ATAATATGCTTTTATTCTTTAAATAGCAAAACCAATTAARAAGA 149359
;
; RESULT 11
; US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-
; Patent No. 6797466
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
;
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
```


;; TITLE OF INVENTION: OF PITUITARY TUMOR TRANSFORMING GENE (PTTG)1 USING PTTG2

;; FILE REFERENCE: 18810-81401
;; CURRENT APPLICATION NUMBER: US/09/854,326
;; CURRENT FILING DATE: 2001-05-11
;; PRIOR APPLICATION NUMBER: US09/730,469
;; PRIOR FILING DATE: 2000-12-04
;; PRIOR APPLICATION NUMBER: US 09/687,911
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 09/569,956
;; PRIOR FILING DATE: 2000-05-12
;; PRIOR APPLICATION NUMBER: US 08/894,251
;; PRIOR FILING DATE: 1999-07-23
;; PRIOR APPLICATION NUMBER: PCT/US86/21463
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: US 60/031,338
;; PRIOR FILING DATE: 1996-11-21
;; NUMBER OF SEQ ID NOS: 68
;; SOFTWARE: Fast-Seq for Windows Version 4.0
;; SEQ ID NO 65
;; LENGTH: 1305
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-854-326-65

Query Match 5.3%; Score 40; DB 3; Length 1305;
Best Local Similarity 53.1%; Pred. No. 0.41; Mismatches 0; Indels 0; Gaps 0;
Matches 85; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 246 CTTATACAAAGAGATTTTATCGCATGTGCGATAAATTAAGGTCAATTAACAAGAA 305
DB 929 CTCGACATAGATTTAAATTTCTAGTGTCTTAGATTTGTGTATATTTCTTAATA 988
QY 306 AACTGAAACGATTTAATTAAGCAAAACATGCTTTCTTAAATCTTAGAAGAGTTTGG 365
DB 989 AAGCATTTATTTGTTAAACAGAAAAAGATATATATCTTAAATCTCTAAATAAATAACCA 1048
QY 366 AGAATCGATGATCAAGAGTGAAGAAATCTGCGATGA 405
DB 1049 TTAAGGAAAAACAGGAGTTATATACTAATAAGGAACAA 1088

RESULT 13
US-09-248-796A-878
; Sequence 878, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 878
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (28)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown

US-09-248-796A-878
Query Match 5.2%; Score 39.8; DB 3; Length 795;
Best Local Similarity 47.2%; Pred. No. 0.39;
Matches 159; Conservative 0; Mismatches 172; Indels 6; Gaps 1;
QY 41 AATGGAATCTAGCATTTATTTGATTTGTTGAGTGTCTTTTGTGTAAGACGGCG 100

DB 254 AAGTTTATTATCTGATTTGATATTGTTATTAGAACCTGTTGAATCACAATGGAAGCAT 313
QY 101 AAAAAAGACACAATGAAAACTGACCACTCCATAGATAAAGAGCATGCGGATATT 160
DB 314 TAGCAAGATTGAATCTCTAAACCTTGAGAAAAAGAAAGAGGTGATGATGATA 373
QY 161 ACGCTAAATACGCAGAAAAAGAAATCGCTGAAGAGTTGCAATATCTATGGGAGCAATAGTTTG 220
DB 374 ACAATAAAATGGTATACCAACCATGAAGAAATTAATTAATAAATCACTAATGTTGTT 433
QY 221 CGAGTTTCATTAAGCGGCAAGAGGTCTTATACAAAGAGATTTTATCGCATGTTGGGATA 280
DB 434 CGATTTTAAATTAATGATATAACAAATTTGATATATATCACTA-----GAGAAG 487
QY 281 AATTAAAGTCAATTAACAACAGAAACTGAAACGACTTTAATTGACAAAAACATGCTTT 340
DB 488 AATTGACGAGGAAATTTTCAAAACAATACTGGACAAAGTTTAAATGTTGAACGAGGTCTTA 547
QY 341 CTAATAATCTTAGAAAGAGTTTGAAGAAATGCGATGA 377
DB 548 AACGATCAAGAGAGAAAGTGTGATGACGAAGATGATGA 584

RESULT 14
US-09-513-999C-14352/C
; Sequence 14352, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14352
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 392
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-14352

Query Match 5.2%; Score 39.2; DB 3; Length 525;
Best Local Similarity 55.0%; Pred. No. 0.49;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 8 ACAATATCATAGAGACTTGGAAATTTTAAAGCAATTTGAATCTAGTATGATTTGATT 67
DB 258 AAAAAAGAGAGAGAAATGGCAAGAGAAAGTTTTTCAAAATTTCTTTCTTTTAAAT 199
QY 68 TGTGAGTGCTTGTGTTTGTGTAAGACGCGGAAAAAGACACAATGAAAACTGACCA 127
DB 198 TAGATTGATTTTATTTTGAACAGACTGGCCCAATGCCACAAGAAATTCCTGTCA 139
QY 128 GCTCCATAGATACAAAGG 147
DB 138 GCACCACCGATGTCCAAAGG 119

RESULT 15
US-09-949-016-17780/C
; Sequence 17780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17780

Query Match 5.2%; Score 39.2; DB 3; Length 601;
Best Local Similarity 55.0%; Pred. No. 0.52;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
Qy 8 ACAATATGATAGACTTGGAAATTTTAAAGCAATTGGAATCTAGTGAATTTATTGGATT 67
Db 181 AAAAAGAAGAGAGAAATGGCAAGAGAAAGTTTTTCAAATTTCTTTCTTTTAAATT 122
Qy 68 TGTGTGAGTGTCTTTTGTGTAAGACGGCGAAAAAGACACAAATGNAAACTGACCA 127
Db 121 TAGATTGAGTTCAATTTATTTGAACAGACTGGGCCAATGTCCACAAAGAAATTCCTGGTCA 62
Qy 128 GCTCCATAGAAATACAAAAGG 147
Db 61 GCACCCCGATGTCCAAAGG 42

Search completed: August 3, 2006, 09:18:19
Job time : 207 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2006, 09:18:35 ; Search time 1473 Seconds
(without alignments)
6331.511 Million cell updates/sec

Title: US-09-732-091-3

Perfect score: 759

Sequence: 1 atggcatacaaatgatag.....cgttgcaaatagaatccatt 759

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759	100.0	759	3	US-09-732-091-3
2	759	100.0	759	8	US-10-433-970-3
3	759	100.0	795	3	US-09-732-091-43
4	759	100.0	795	8	US-10-433-970-43
5	727	95.8	762	8	US-10-335-977-4400
6	727	95.8	762	8	US-10-335-977-4401
7	727	95.8	768	8	US-10-335-977-4402
8	717.4	94.5	762	8	US-10-433-970-47
9	717.4	94.5	798	8	US-10-433-970-45
10	365.8	48.2	744	8	US-10-335-977-2937
11	359.4	47.4	714	8	US-10-335-977-2936
12	271.4	35.8	547	3	US-09-882-227-413
13	153	20.2	153	3	US-09-732-091-36
14	153	20.2	153	8	US-10-433-970-36
15	114	15.0	114	3	US-09-732-091-33
16	114	15.0	114	8	US-10-433-970-33
17	108	14.2	108	3	US-09-732-091-35

18	108	14.2	108	8	US-10-433-970-35	Sequence 35, Appl	
19	90	11.9	90	3	US-09-732-091-32	Sequence 32, Appl	
20	90	11.9	90	3	US-09-732-091-34	Sequence 34, Appl	
21	90	11.9	90	8	US-10-433-970-32	Sequence 32, Appl	
22	90	11.9	90	8	US-10-433-970-34	Sequence 34, Appl	
23	47.6	6.3	1214	8	US-10-424-599-102083	Sequence 102083,	
24	47.6	6.3	1243	9	US-10-425-115-17217	Sequence 17217,	
c	25	47	6.2	627	8	US-10-021-323-9336	Sequence 9336, Ap
26	46.2	6.1	476	9	US-10-425-115-50533	Sequence 50533, A	
27	46.2	6.1	3588	8	US-10-282-122A-16553	Sequence 16553, A	
c	28	45.4	6.0	547	4	US-09-925-065A-115849	Sequence 115849,
c	29	45.4	6.0	547	5	US-09-925-065A-115849	Sequence 115849,
c	30	45.4	6.0	548	12	US-10-301-480-214754	Sequence 214754,
c	31	45.4	6.0	548	12	US-10-301-480-214754	Sequence 214754,
c	32	45.4	6.0	1062	9	US-10-425-115-120013	Sequence 120013,
c	33	45.2	6.0	564	4	US-09-925-065A-109807	Sequence 109807,
34	45.2	6.0	564	4	US-09-925-065A-109808	Sequence 109808,	
35	45.2	6.0	564	4	US-09-925-065A-109809	Sequence 109809,	
36	45.2	6.0	564	5	US-09-925-065A-109807	Sequence 109807,	
37	45.2	6.0	564	5	US-09-925-065A-109808	Sequence 109808,	
38	45.2	6.0	564	5	US-09-925-065A-109809	Sequence 109809,	
39	45.2	6.0	587	12	US-10-301-480-209418	Sequence 209418,	
40	45.2	6.0	587	12	US-10-301-480-209419	Sequence 209419,	
41	45.2	6.0	587	12	US-10-301-480-822827	Sequence 822827,	
42	45.2	6.0	587	12	US-10-301-480-822827	Sequence 822827,	
43	45.2	6.0	587	12	US-10-301-480-822828	Sequence 822828,	
44	45.2	6.0	587	12	US-10-301-480-822829	Sequence 822829,	
45	45.2	6.0	3272	10	US-10-750-185-62663	Sequence 62663, A	

ALIGNMENTS

RESULT 1

US-09-732-091-3
; Sequence 3, Application US/09732091
; Patent No. US20020107368A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-09-732-091-3

Query Match	100.0%	Score 759;	DB 3;	Length 759;
Best Local Similarity	100.0%;	Pred. No. 7.8e-179;	Mismatches 0;	Indels 0;
Matches 759;	Conservative 0;			
QY	1	ATGGCATACAAAATATGATAGAGACTTTGGAAATTTTAAAGCAATTTGGAACTTAGTCAATTA	60	
Db	1	ATGGCATACAAAATATGATAGAGACTTTGGAAATTTTAAAGCAATTTGGAACTTAGTCAATTA	60	
QY	61	TTGGATTTTGTGAGTGTCTTTGTTTGTAAAGACGGCGGCAAAAGACACAAATGAAAAA	120	
Db	61	TTGGATTTTGTGAGTGTCTTTGTTTGTAAAGACGGCGGCAAAAGACACAAATGAAAAA	120	
QY	121	CTGACGAGCTCCATAGAAATACAAAGCGATGCGGATGATTACGCTAAATACGCAGAGAAGA	180	
Db	121	CTGACGAGCTCCATAGAAATACAAAGCGATGCGGATGATTACGCTAAATACGCAGAGAAGA	180	
QY	181	ATCGCTGAAGAGTGTGCAATACTATGCGAGCAATAGTTTTTCGAGTTTCATTAAAGCGAA	240	
Db	181	ATCGCTGAAGAGTGTGCAATACTATGCGAGCAATAGTTTTTCGAGTTTCATTAAAGCGAA	240	

QY 241 GGAGTCTTATACAAAGAGATTTTATGCGATGCTGCGATGATTAAGGTCAATTACAAC 300
DB 241 GGAGTCTTATACAAAGAGATTTTATGCGATGCTGCGATGATTAAGGTCAATTACAAC 300
QY 301 AAGAAACTGAAACGACTTTAAATGGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 360
DB 301 AAGAAACTGAAACGACTTTAAATGGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 360
QY 361 TTGGAAGAAATGGAATGATGAAGAGTGAAGAAATGTCGATGAATATCCATAAAAAC 420
DB 361 TTGGAAGAAATGGAATGATGAAGAGTGAAGAAATGTCGATGAATATCCATAAAAAC 420
QY 421 ACGGCAATTTAAACAGACAGCCTTAAGCGCGGACTTTAAACGCTTTTAAATGGG 480
DB 421 ACGGCAATTTAAACAGACAGCCTTAAGCGCGGACTTTAAACGCTTTTAAATGGG 480
QY 481 GGTTTTAAATCTTATCAATAGCTGCTCAATGTCGATGCTGCGATGAATATCCATAAAAAC 540
DB 481 GGTTTTAAATCTTATCAATAGCTGCTCAATGTCGATGCTGCGATGAATATCCATAAAAAC 540
QY 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTACAAGAACTCTGAGCTTTTAAACA 600
DB 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTACAAGAACTCTGAGCTTTTAAACA 600
QY 601 GGTCTGTTGGCTGGATCATTAAGGCGTATGGAAGGATGATATGAGGCGCGCT 660
DB 601 GGTCTGTTGGCTGGATCATTAAGGCGTATGGAAGGATGATATGAGGCGCGCT 660
QY 721 GCCAATGGAGATGAAGAGTCTGTCGAATAGAATCCATT 759
DB 721 GCCAATGGAGATGAAGAGTCTGTCGAATAGAATCCATT 759

RESULT 2

US-10-433-970-3
; Sequence 3, Application US/10433970
; Publication No. US20040138415A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-10-433-970-3

Query Match 100.0%; Score 759; DB 8; Length 759;
Best Local Similarity 100.0%; Pred. No. 7.8e-179;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCATACAATATGATAGAGACTTGGAAATTTTAAAGCAATGGAATCTAGTGATTTA 60
DB 1 ATGGCATACAATATGATAGAGACTTGGAAATTTTAAAGCAATGGAATCTAGTGATTTA 60
QY 61 TTGGATTTGTTGAGTCTGTTTTCGTAAGACGCGGAAAAAGACAAATGAAAAA 120
DB 61 TTGGATTTGTTGAGTCTGTTTTCGTAAGACGCGGAAAAAGACAAATGAAAAA 120

QY 121 CTGACCGCTCCATAGATAACAAAGGCGATGCGGATGATTAAGCTAAATACGAGAAAGA 180
DB 121 CTGACCGCTCCATAGATAACAAAGGCGATGCGGATGATTAAGCTAAATACGAGAAAGA 180
QY 181 ATCGCTGAAGAGTTGCAATCTATCGGAGCAATAGTTTTGCGAGTTTCAATTAAGGCGAA 240
DB 181 ATCGCTGAAGAGTTGCAATCTATCGGAGCAATAGTTTTGCGAGTTTCAATTAAGGCGAA 240
QY 241 GGAGTCTTATCAAAAGAGATTTTATGCGATGCTGCGATGAATTAAGGTCAATTACAAC 300
DB 241 GGAGTCTTATCAAAAGAGATTTTATGCGATGCTGCGATGAATTAAGGTCAATTACAAC 300
QY 301 AAGAAACTGAAACGACTTTAAATGGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 360
DB 301 AAGAAACTGAAACGACTTTAAATGGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 360
QY 361 TTGGAAGAAATGGAATGATGAAGAGTGAAGAAATGTCGATGAATATCCATAAAAAC 420
DB 361 TTGGAAGAAATGGAATGATGAAGAGTGAAGAAATGTCGATGAATATCCATAAAAAC 420
QY 421 ACGGCAATTTAAACAGACAGCCTTAAGCGCGGACTTTAAACGCTTTTAAATGGG 480
DB 421 ACGGCAATTTAAACAGACAGCCTTAAGCGCGGACTTTAAACGCTTTTAAATGGG 480
QY 481 GGTTTTAAATCTTATCAATAGCTGCTCAATGTCGATGCTGCGATGAATATCCATAAAAAC 540
DB 481 GGTTTTAAATCTTATCAATAGCTGCTCAATGTCGATGCTGCGATGAATATCCATAAAAAC 540
QY 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTACAAGAACTCTGAGCTTTTAAACA 600
DB 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTACAAGAACTCTGAGCTTTTAAACA 600
QY 601 GGTCTGTTGGCTGGATCATTAAGGCGTATGGAAGGATGATATGAGGCGCGCT 660
DB 601 GGTCTGTTGGCTGGATCATTAAGGCGTATGGAAGGATGATATGAGGCGCGCT 660
QY 661 TATAGGGTAACCATACCGGCGATGCTGTCGAATAGAATCCATT 759
DB 661 TATAGGGTAACCATACCGGCGATGCTGTCGAATAGAATCCATT 759
QY 721 GCCAATGGAGATGAAGAGTCTGTCGAATAGAATCCATT 759
DB 721 GCCAATGGAGATGAAGAGTCTGTCGAATAGAATCCATT 759

RESULT 3

US-09-732-091-43
; Sequence 43, Application US/09732091
; Patent No. US20020107368A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 43
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-09-732-091-43

Query Match 100.0%; Score 759; DB 3; Length 795;
Best Local Similarity 100.0%; Pred. No. 8e-179;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCATACAATATGATAGAGACTTGGAAATTTTAAAGCAATGGAATCTAGTGATTTA 60
DB 37 ATGGCATACAATATGATAGAGACTTGGAAATTTTAAAGCAATGGAATCTAGTGATTTA 96

```
Qy 61 TTGGAATTTGTTGAGTGTCTGTTTGTGTAAGAAGCGGCAAAAAGACACAATGAAAAA 120
Db 97 TTGGAATTTGTTGAGTGTCTGTTTGTGTAAGAAGCGGCAAAAAGACACAATGAAAAA 156
Qy 121 CTGACGAGCTCCATAGATAAACAAGGCGATGCGGATGATTAGCTTAATACGCAAGAA 180
Db 157 CTGACGAGCTCCATAGATAAACAAGGCGATGCGGATGATTAGCTTAATACGCAAGAA 216
Qy 181 ATCCCTGAAGAGTTGCAATCTATGGGAGCAATAGTTTGGGAGTTTCATTAAAGCGCAA 240
Db 217 ATCCCTGAAGAGTTGCAATCTATGGGAGCAATAGTTTGGGAGTTTCATTAAAGCGCAA 276
Qy 241 GGAGTCTTATACAAAGAGATTTATGCGATGTGCGGATGATAAATTAAGGTCAATTAACA 300
Db 277 GGAGTCTTATACAAAGAGATTTATGCGATGTGCGGATGATAAATTAAGGTCAATTAACA 336
Qy 301 AAGAAACTGAAACGACTTTTAAATGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 360
Db 337 AAGAAACTGAAACGACTTTTAAATGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 396
Qy 361 TTGGAAGAAATGGATGATGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGT 420
Db 397 TTGGAAGAAATGGATGATGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGT 456
Qy 421 ACGGACAAATTTAAACAGACAGCTTAAAGCGCGGCGACTTTAAAGCGCTGTTTAAATGGG 480
Db 457 ACGGACAAATTTAAACAGACAGCTTAAAGCGCGGCGACTTTAAAGCGCTGTTTAAATGGG 516
Qy 481 GGTGTTTAAATCTTATCAATTAGCTGTCTATTGTCGAAATGCGGTGCGCAAAACCAATCTTA 540
Db 517 GGTGTTTAAATCTTATCAATTAGCTGTCTATTGTCGAAATGCGGTGCGCAAAACCAATCTTA 576
Qy 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAAACA 600
Db 577 GGGCGTGGTTTATCGCTTGGCGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAAACA 636
Qy 601 GGTCTCTGTGGCTGATCATTTACAGCGCTATGACAGCGATTGATATGACAGGCGCGCT 660
Db 637 GGTCTCTGTGGCTGATCATTTACAGCGCTATGACAGCGATTGATATGACAGGCGCGCT 696
Qy 721 GCCAATGGAGATAAGAAGTCGTTGCCAATAGAATCCATT 759
Db 757 GCCAATGGAGATAAGAAGTCGTTGCCAATAGAATCCATT 795
```

RESULT 4

```
US-10-433-970-43
; Sequence 43, Application US/10433970
; Publication No. US20040138415A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
```

```
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; TITLE OF INVENTION: THEREOF
```

```
; FILE REFERENCE: 7969-091-999
```

```
; CURRENT APPLICATION NUMBER: US/10/433,970
```

```
; PRIOR FILING DATE: 2003-06-06
```

```
; PRIOR APPLICATION NUMBER: 09/732,091
```

```
; PRIOR FILING DATE: 2000-12-07
```

```
; NUMBER OF SEQ ID NOS: 48
```

```
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 43
```

```
; LENGTH: 795
```

```
; TYPE: DNA
```

```
; ORGANISM: Helicobacter sp.
```

```
US-10-433-970-43
```

```
Query Match 100.0%; Score 759; DB 8; Length 795;
Best Local Similarity 100.0%; Pred. No. 8e-179;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGGCATACAAATATGATAGAGACTTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
Db 37 ATGGCATACAAATATGATAGAGACTTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 96
Qy 61 TTGGATTTTGTGAGGTGCTTGTGTTTGGTAAAGACGGCGGAAAAAGACACAATGAAAAA 120
Db 97 TTGGATTTTGTGAGGTGCTTGTGTTTGGTAAAGACGGCGGAAAAAGACACAATGAAAAA 156
Qy 121 CTGACAGCTCCATAGATAAACAAGGCGATGCGGATGATTAGCTTAATACGCAAGAA 180
Db 157 CTGACAGCTCCATAGATAAACAAGGCGATGCGGATGATTAGCTTAATACGCAAGAA 216
Qy 181 ATCCCTGAAGAGTTGCAATCTATGGGAGCAATAGTTTGGGAGTTTCATTAAAGCGCAA 240
Db 217 ATCCCTGAAGAGTTGCAATCTATGGGAGCAATAGTTTGGGAGTTTCATTAAAGCGCAA 276
Qy 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGGATGATAAATTAAGGTCAATTAACA 300
Db 277 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGGATGATAAATTAAGGTCAATTAACA 336
Qy 301 AAGAAACTGAAACGACTTTTAAATGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 360
Db 337 AAGAAACTGAAACGACTTTTAAATGAACAAACATGCTTTCTAAATCTTAGAAAGAGT 396
Qy 361 TTGGAAGAAATGGATGATGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGT 420
Db 397 TTGGAAGAAATGGATGATGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGT 456
Qy 421 ACGGACAAATTTAAACAGACAGCTTAAAGCGCGGCGACTTTAAAGCGCTGTTTAAATGGG 480
Db 457 ACGGACAAATTTAAACAGACAGCTTAAAGCGCGGCGACTTTAAAGCGCTGTTTAAATGGG 516
Qy 481 GGTGTTTAAATCTTATCAATTAGCTGTCTATTGTCGAAATGCGGTGCGCAAAACCAATCTTA 540
Db 517 GGTGTTTAAATCTTATCAATTAGCTGTCTATTGTCGAAATGCGGTGCGCAAAACCAATCTTA 576
Qy 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAAACA 600
Db 577 GGGCGTGGTTTATCGCTTGGCGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAAACA 636
Qy 601 GGTCTCTGTGGCTGATCATTTACAGCGCTATGACAGCGATTGATATGACAGGCGCGCT 660
Db 637 GGTCTCTGTGGCTGATCATTTACAGCGCTATGACAGCGATTGATATGACAGGCGCGCT 696
Qy 661 TATAGGTTAACCATACCGGCGATGCTGTTGGTGGCCACTTTACGCTTAAAAACACAGCAA 720
Db 697 TATAGGTTAACCATACCGGCGATGCTGTTGGTGGCCACTTTACGCTTAAAAACACAGCAA 756
Qy 721 GCCAATGGAGATAAGAAGTCGTTGCCAATAGAATCCATT 759
Db 757 GCCAATGGAGATAAGAAGTCGTTGCCAATAGAATCCATT 795
```

RESULT 5

```
US-10-335-977-4400
```

```
; Sequence 4400, Application US/10335977
```

```
; Publication No. US20040052799A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: DOUGLAS SMITH et al
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
```

```
; RELATING TO HELICOBACTER PYLORI FOR
```

```
; DIAGNOSTICS AND THERAPEUTICS
```

```
; NUMBER OF SEQUENCES: 10031
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: LAHIVE & COCKFIELD
```

```
; STREET: 28 State Street
```

```
; CITY: Boston
```

```
; STATE: Massachusetts
```

```
; COUNTRY: USA
```


Best Local Similarity 97.4%; Pred. No. 7.6e-171;
Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
DB 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60

QY 61 TTGGATTTGTTGAGGTGCTGTTTTGTTGTAAGACGGCGAAAGACACAAATGAAAAA 120
DB 61 TTGGATTTGTTGAGGTGCTGTTTTGTTGTAAGACGGCGAAAGACACAAATGAAAAA 120

QY 121 CTGACAGCTCCATAGATACAAAGGCGATGCGGATGATCGCTAAATACGAGAAAGA 180
DB 121 CTCACAAGCTCCATAGATACAAAGGCGATGCGGATGATCGCTAAATACGAGAAAGA 180

QY 181 ATCGCTGAAGAGTTCGAATCTATGCGAGCAATAGTTTTGCGAGTTTCAATTAAGCGGAA 240
DB 181 ATCGCTGAAGAGTTCGAATCTATGCGAGCAATAGTTTTGCGAGTTTCAATTAAGCGGAA 240

QY 241 GGAGTCTTATACAAAGAGATTTTATGCGATGCTGCGATAAATTAAGGTCAATTACAC 300
DB 241 GGAGTCTTATACAAAGAGATTTTATGCGATGCTGCGATAAATTAAGGTCAATTACAC 300

QY 301 AAGAAAACCTGAACGACTTTAATTGAAACAAACATGCTTTCTTAAATCTTGAAGAGAGT 360
DB 301 AAGAAAACCTGAACGACTTTAATTGAAACAAACATGCTTTCTTAAATCTTGAAGAGAGC 360

QY 361 TTGGAGAAATGGATGATGAGAGAGTGAAGAAATGTGCGATGAATTTCCATTAAGAAC 420
DB 361 CTAGAAGAAATGGATGATGAGAGAGTGAAGAAATGTGCGATGAATTTCCATTAAGAAC 420

QY 421 ACGGCAATTTTAAACAGACAAAGCCTTAAGCGCGGACCTTAAACGCTGTTTAAATGGG 480
DB 421 ACGGCAATTTTAAACAGACAAAGCCTTAAGCGCGGACCTTAAACGCTGTTTAAATGGGA 480

QY 481 GGTTTTAAATCTTATCAATAGTCTGATGTTGCGAATCGGTCGCAAAACCAATTTCTA 540
DB 481 GGTTTTAAATCTTATCAATAGTCTGATGTTGCGAATCGGTCGCAAAACCAATTTCTA 540

QY 541 GGGCGTGTTTATCGCTGGCGGCAATCAGGTGCTTACAGAACTCTGAGCTTTTAAACA 600
DB 541 GGGCGTGTTTATCGCTGGCGGCAATCAGGTGCTTACAGAACTCTGAGCTTTTAAACA 600

QY 601 GGTCTGTTGCTGATCATTTACAGCGGTATGAGCAGCGATGATATTGCGGGCGGCT 660
DB 601 GCGCTGTTGCTGATCATTTACAGCGGTATGAGCAGCGATGATATTGCGGGCGGCT 660

QY 661 TATAGGGTAACCATACCGGATGCAATTTGTTGCGATGCTTACGCTTAAACACAGCAA 720
DB 661 TATAGGGTAACCATACCGGATGCAATTTGTTGCGATGCTTACGCTTAAACACGCAA 720

QY 721 GCCAATGAGATAGAAGTCTGTTGCAATAGAAATCCATT 759
DB 721 GCCAATGAGATAGAAGTCTGTTGCAATAGAAATCCGTT 759

RESULT 7

US-10-335-977-4402
; Sequence 4402, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 4402:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...768
SEQUENCE DESCRIPTION: SEQ ID NO: 4402:
US-10-335-977-4402

Query Match 95.8%; Score 727; DB 8; Length 768;
Best Local Similarity 97.4%; Pred. No. 7.6e-171;
Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
DB 10 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATCTGGAATCTAGTGATTTA 69

QY 61 TTGGATTTGTTGAGGTGCTGTTTTGTTGTAAGACGGCGAAAGACACAAATGAAAAA 120
DB 70 TTGGATTTGTTGAGGTGCTGTTTTGTTGTAAGACGGCGAAAGACACAAATGAAAAA 129

QY 121 CTGACAGCTCCATAGATACAAAGGCGATGCGGATGATCGCTAAATACGAGAAAGA 180
DB 130 CTCACAAGCTCCATAGATACAAAGGCGATGCGGATGATCGCTAAATACGAGAAAGA 189

QY 181 ATCGCTGAAGAGTTCGAATCTATGCGGAGCAATAGTTTTGCGAGTTTCAATTAAGCGGAA 240
DB 190 ATCGCTGAAGAGTTCGAATCTATGCGGAGCAATAGTTTTGCGAGTTTCAATTAAGCGTAA 249

QY 241 CGAGTCTTATACAAAGAGATTTTATGCGATGCTGCGATAAATTAAGGTCAATTAACAAC 300
DB 250 GGAGTCTTATACAAAGAGATTTTATGCGATGCTGCGATAAATTAAGGTCAATTAACAAC 309

QY 301 AAGAAAACCTGAACGACTTTAATTGAAACAAACATGCTTTCTTAAATCTTGAAGAGAGC 360
DB 310 AAGAAAACCTGAACGACTTTAATTGAAACAAACATGCTTTCTTAAATCTTGAAGAGAGC 369

QY 361 TTGGAAGAAATGGATGATGAGAGAGTGAAGAAATGTGCGATGAATTTATCATTAAGAAC 420
DB 370 CTAGAAGAAATGGATGATGAGAGAGTGAAGAAATGTGCGATGAATTTATCATTAAGAAC 429

QY 421 ACGGCAATTTTAAACAGACAGCCTTAAAGCGGCGACCTTTAAGCGTGTTTAAATGGG 480
DB 430 ACGGCAATTTTAAACAGACAGCCTTAAAGCGGCGACCTTTAAGCGTGTTTAAATGGGA 489

QY 481 GGTTTTAAATCTTATCAATAGTCTGATGTTGCGAATCGGTCGCAAAACCAATTTCTA 540

Db 490 GGCTTTAAATCTTATCAATAGCTGTCTATGTTGCGAATCGGTTGCAAAAACCAATCTTA 549
Qy 541 GGGCTGTTTATCGCTTGGGCAATCAGTGTCTTACAGAACTCTGAGCTTTTAAACA 600
Db 550 GGGCTGTTTATCGCTTGGGCAATCAAGTGTCTTACAGAACTCTGAGCTTTTAAACA 609
Qy 601 GGTCTGTTGGCTGGATCATTTACAGGCGTATGACAGCGATTGATATTGACAGGCGCGCT 660
Db 610 GGGCTGTTGGCTGGATCATTTACAGGCGTATGACAGCGATTGATATTGACAGGCGCGCT 669
Qy 661 TATAGGTAACCATACCGGATGATTTGTTGCTTCCACTTTACGCCCTAAACCAACAGCAA 720
Db 670 TATAGGTAACCATACCGGATGATTTGTTGCTTCCACTTTACGCCCTAAACCAACAGCAA 729
Qy 721 GCCAATGAGATAAGAGTGTGTTGCAATAGATCCATT 759
Db 730 GCCAATGAGATAAGAGTGTGTTGCAATAGATCCGTT 768

RESULT 8

US-10-433-970-47
; Sequence 47, Application US/10433970
; Publication No. US20040138415A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; PRIOR FILING DATE: 2003-06-06
; PRIOR FILING DATE: 2003-06-06
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Helicobacter sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(762)
US-10-433-970-47

Query Match 94.5%; Score 717.4; DB 8; Length 762;
Best Local Similarity 96.6%; Pred. No. 1.9e-168;
Matches 733; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTGGAAATCTAGTGATTTA 60
Db 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTGGAAATCTAGTGATTTA 60
Qy 61 TTGGATTTGTTGAGGTGCTTTTGGTAAAGCGCGGAAAGACACAAATGAAAA 120
Db 61 TTGGATTTGTTGAGGTGCTTTTGGTAAAGCGCGGAAAGACACAAATGAAAA 120
Qy 121 CTGACCAAGCTCCATAGATAACAAAGGATGCGCATGATTAACGTTAAATACGCAAGAGA 180
Db 121 CTGACCAAGCTCCATAGATAACAAAGGATGCGCATGATTAACGTTAAATACGCAAGAGG 180
Qy 181 ATCGCTGAAGATTGCAATCTATGGGAGCAATAGTTTGGAGTTTCATTAAGGCGAA 240
Db 181 ATCGCTGAAGATTGCAATCTATGGGAGCAATAGTTTGGAGTTTCATTAAGGCGAA 240
Qy 241 GGAGTCTTATCAAGAGATTTTATGCGATGTCGCGATAATTAAGGTCATTAACAC 300
Db 241 GGAGTCTTATCAAGAGATTTTATGCGATGTCGCGATAATTAAGGTCATTAACAC 300
Qy 301 AAGAAAACTGAACGACTTTAAATGAACAAACATGCTTTCTAAATCTTGAAGAAGT 360
Db 301 AAGAAAACTGAACGACTTTGAATGAACAAACATGCTTTCTAAATCTTGAAGAAGT 360

Qy 361 TTGGAAGAAATGGATGATGAAGAACTGAAGAAATGTGCGATGAATATCCATAAAAAAC 420
Db 361 TTGGAAGAAATGGATGATGAAGAACTGAAGAAATGTGCGATGAATATCCATAAAAAAC 420
Qy 421 ACAGCAATTTAAACAGACAAGCCTTAAGCGCGGCGAATTTAAACGCTGTTTAAATGGGG 480
Db 421 ACTGACAATTTGAACAGACAAGCCTTTAAGCGCGGCGACTTTAAACGCTGTTTAAATGGGA 480
Qy 481 GGTTTTAAATCTTATCAATTAAGTCTGCTTGTGCGAATGCGGTGCGCAAAAACCAATCTTA 540
Db 481 GGTTTTAAATCTTATCAATTAAGTCTGCTTGTGCGAATGCGGTGCGCAAAAACCAATCTTA 540
Qy 541 GGGCTGTTTATCGCTTGGCGCAATCAGTGTCTTCAAGAACTCTGAGCTTTTAAACA 600
Db 541 GGGCTGTTTATCGCTTGGCGCAATCAGTGTCTTCAAGAACTCTGAGCTTTTAAACA 600
Qy 601 GGTCTGTTGGCTGGATCATTTACAGCGGTATGGAAGCGATTTGATATGAGGCGCGGCT 660
Db 601 GGTCTGTTGGCTGGATCATTTACAGCGGTATGGAAGCGATTTGATATGAGGCGCGGCT 660
Qy 661 TATAGGTAACCATACCGGATGATTTGTTGCGAATGCGGTGCGCAAAAACCAACAGCAA 720
Db 661 TATAGGTAACCATACCGGATGATTTGTTGCGAATGCGGTGCGCAAAAACCAACAGCAA 720
Qy 721 GCCAATGAGATAAGAGTGTGTTGCAATAGATCCATT 759
Db 721 GCCAATGAGATAAGAGTGTGTTGCAATAGATCCATT 759

RESULT 9

US-10-433-970-45
; Sequence 45, Application US/10433970
; Publication No. US20040138415A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; PRIOR FILING DATE: 2003-06-06
; PRIOR FILING DATE: 2003-06-06
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Helicobacter sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(798)
US-10-433-970-45

Query Match 94.5%; Score 717.4; DB 8; Length 798;
Best Local Similarity 96.6%; Pred. No. 1.9e-168;
Matches 733; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTGGAAATCTAGTGATTTA 60
Db 37 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTGGAAATCTAGTGATTTA 96
Qy 61 TTGGATTTGTTGAGGTGCTTTTGGTAAAGCGCGGAAAGACACAAATGAAAA 120
Db 97 TTGGATTTGTTGAGGTGCTTTTGGTAAAGCGCGGAAAGACACAAATGAAAA 156
Qy 121 CTGACCAAGCTCCATAGATAACAAAGGATGCGCATGATTAACGTTAAATACGCAAGAGA 180
Db 157 CTCAAGCTCCCTAGATAACAAAGGATGCGCATGATTAACGTTAAATACGCAAGAGG 216
Qy 181 ATCGCTGAAGATTGCAATCTATGGGAGCAATAGTTTGGAGTTTCTAATGAAGCGAA 240

Db 217 ATCGCTGAAGAGTTGCCAATCTATCGGAGCAATAGTTTTGGAGTTTCATTAAAGCGGAA 276
Qy 241 GGAGCTTTATACAAAGAGATTTTATCGCATGTGTGCATATAAATTAAGGTCAATTACAAAC 300
Db 277 GGGGTCTTATACAAAGAGATCTATCGCATGTGTGGGATAAATTAAGGTCAATTACAAAC 336
Qy 301 AAGAAACTGAACCGACTTTTAAATGGAACAAACATCGCTTTCTAAATCTTGAAGAAGT 360
Db 337 AAGAAACTGAACCGACTTTTAAATGGAACAAACATCGCTTTCTAAATCTTGAAGAAGT 396
Qy 361 TTGGAAGAAATGGATGATGAAGAGTGAAGAAATGGGATGAATATCCATATAAAG 420
Db 397 TTAGAGGAATGGATGATGAAGAGTGAAGAAATGGGATGAATATCCATATAAAG 456
Qy 421 ACGGCAATTTAAACAGACAAAGCCTTAAGCGCGCGACTTTAAAGCTGTTTAAATGGG 480
Db 457 ACTGACAAATTTGAACAGACAGCTTTTAAAGCGCGCGACTTTAAAGCTGTTTAAATGGG 516
Qy 481 GGTTTTAAATCTTATCAATAGCTGTCTATGTTGCGAATGCGGTGCGAAACCACTTCTA 540
Db 517 GGCTTTAAATCTTATCAATAGCTGTCTATGTTGCGAATGCGGTGCGAAACCACTTCTA 576
Qy 541 GGGCGTGTGTTATCGCTGCGGCGCAATCAGGTGCTTACAGAACTCTGAGCTTTTAAACA 600
Db 577 GGGCGTGTGTTATCGCTGCGGCGCAATCAGGTGCTTACAGAACTCTGAGCTTTTAAACA 636
Qy 601 GGTCTGTGTGCTGATCATTTACAGGCGCTATGAGCGATGATGATTCGAGGCGCGCT 660
Db 637 GGCCCTGTGTGCTGATCATTTACAGGCGCTATGAGCGATGATGATTCGAGGCGCGCT 696
Qy 661 TATAGGGTAAACATACCGGCGATGATGTTGTTGCGAATGCGGTGCGAAACCACTTCTA 720
Db 697 TATAGGGTAAACATACCGGCGATGATGTTGTTGCGAATGCGGTGCGAAACCACTTCTA 756
Qy 721 GCCAATGAGATAGAAGTCTGTTGCAATAGCAATCCATT 759
Db 757 GCCAATGAGATAGAAGTCTGTTGCAATAGCAATCCATT 795

RESULT 10

US-10-335-977-2937
; Sequence 2937, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 2937:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 744 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...744
; SEQUENCE DESCRIPTION: SEQ ID NO: 2937:
US-10-335-977-2937

Query Match 48.2%; Score 365.8; DB 8; Length 744;
Best Local Similarity 71.5%; Pred. No. 1.1e-80;
Matches 513; Conservative 0; Mismatches 192; Indels 12; Gaps 2;

Qy 17 ATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGAATTTATTTGGATTTGTTGAGG 76
Db 26 ATAGCGACTTGGAAATTTTAAAGCGATTAAAGCTCTAGCGATTTGAAAGATTTGTTGATG 85
Qy 77 TCCTTGTGTTTGGTAAAGACGCGGAAAAAGACACAAATGAAAACTGACCAGCTCCATAG 136
Db 86 CCGCTTGTCTATGATGAAGATGGCACACTAAGAAATGAATGAAGATTTGCAAGTTTACAG 145
Qy 137 AATACAAAAGGATGGCGATGATTACGCTAAATPACGACAGAAAGATCGCTGAAGAGTTGC 196
Db 146 AATACAAAAGGATTTGGCCATGATTACGCCAATATCCCAAGAGATCGCTGAAGAAATGC 205
Qy 197 AATATATGGGAGCAATAGTTTTTGGAGTTTCATTAAGCGGAGAGTCTTTATACAAAG 256
Db 206 AGCGTTATGGGCGCAATAGTTTTTGGCAATTTTTTTAGAGATGAAGGGGTCTTTATACAAAG 265
Qy 257 AGATTTTATGCGATGTCGCGATTAATTAAGTCAATTAACAACAAGAACTGAAACGA 316
Db 266 AGATTTTGTGCGATGTCGCGATCAATTTGGATTAATTAACAATGAAGATCTGCAACCT 325
Qy 317 CTTTAAATGAACAAACACATGCTTTCTAAATCTTGAAGAAAGATTTTGAAGAAATGATG 376
Db 326 CTTTCAATGAGCAAAACATGCTTTCTAAACTTTTGAAGATAGTTTAGAAGAAATGATG 385
Qy 377 ATGAAGAAGTGAAGAAATGTCGATGAATATCCATATAAAGAAACACGACAA----- 428
Db 386 GGAGAGAGATTAAGAACTCTCGCATGGATTGGGCATGCCAATATTGATAAAGTATTG 445
Qy 429 -TTTAAACAGACAGCCTTAAGCGCGGCGACTTTTAACGCTGTTTAAATGCGGGGTTTA 487
Db 446 GTGAAAAACAAACAGTCCCTTAATCGCATCTGTTTTTAACGCTGTTTAAAGCGGCGGCTC 505
Qy 488 AATCTTATCAATTAGCTGTCATTGTTGCGAATGCGGTGCGCAAAACCATTTCTAGGGCGTG 547
Db 506 ATCTTATCGGTTGGCTGTAGCTGTGAGATGCAATGTTAGCAAACTCTAGGGCATG 565
Qy 548 GTTATTCG---CTTGGGGCAATCAGGTGCTTTACAAGAACTCTGAGCTTTTAAACAGGTC 604
Db 566 GTTATTCGCTCTGTGGTGAAGTAGCGCTTTAAAAAACTCTAGACATTTTAGTGGCC 625
Qy 605 CTGTTGGCTGGATCATTTACAGCGGTATGGAACGAGTTGATATTGAGGCGCGGCTTATA 664
Db 626 CTATTGTTGGGTCAATTACGGGCGGTTAGTGAGCATCAATCTTGTGGCGCGGCTTATA 685
Qy 665 GGGTAAACCATACCGGCTGCTGTTGGTTGGCCACTTTAGCCCTTAAACACACAGCAAG 721
Db 686 GGGTGACCGTGCTCGATCGGTTTTTAGTTGGCCACTTACGCAAAAAAATAAAGACAG 742

; Sequence 2936, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 2936:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 714 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...714
; SEQUENCE DESCRIPTION: SEQ ID NO: 2936:

US-10-335-977-2936
Query Match 47.4%; Score 359.4; DB 8; Length 714;
Best Local Similarity 71.4%; Pred. No. 4.3e-79;
Matches 506; Conservative 0; Mismatches 191; Indels 12; Gaps 2;

Qy	25	TTGGAATTTTAAAGCAATTCGAATCTAGTATTATTTGGATTCTTTGAGGTGCTTCTT	84
Db	1	TTGGAATTTTAAAGCGATTAAAGCAATTCGAATCTAGTATTATTTGGATTCTTTGAGGTGCTTCTT	60
Qy	85	TTTGTAAAGACGGCGAAAGACACAAATGAAACCTGACAGCTCCATAGATACAA	144
Db	61	TATGATGAGATGGCAGCACTAAGATGATGAGATTTGACAGTTTACAGATACCA	120
Qy	145	AGGATGCGCATGATTACGCTTAAATACGCAGAAAGATTCGCTGAAGAGTTTCAAT	204
Db	121	AGGATGCGCATGATTACGCTTAAATACGCAGAAAGATTCGCTGAAGAGTTTCAAT	180
Qy	205	GGGAGCAATAGTTTTCGAGTTTCAATTAAGCGAGAGTCTTATACAAAGATTTTA	264
Db	181	GGGAGCAATAGTTTTCGAGTTTCAATTAAGCGAGAGTCTTATACAAAGATTTTG	240
Qy	265	TGGATGTGCGCAATTAAGGTCATTAACAAAGAACTGAAACGACTTTAATT	324
Db	241	TGGATGTGCGCAATTAAGGTCATTAACAAAGAACTGAAACGACTTTTGAAT	300

Qy	325	GAACAAACATGCTTTCTAAATCTTAGAAAGAACTTTGAAAGAAATGATGATGAAGAA	384
Db	301	GAGCAAAACATGCTTTCTAAATCTTAGAAAGAACTTTGAAAGAAATGATGATGAAGAA	360
Qy	385	GTGAAAGAAATGTCGATGAATATCCATAAATAACACGACAA-----TTTAAAC	435
Db	361	ATTAAGAACTCTCGATGGATTGGCATGCGCAATATTTGATAAATGATTGTTGTTGAAAC	420
Qy	436	AGACAAGCCTTAAGCGGCGGACCTTAAGCGCTTTAAATGCGGGGTTTAAATCTTAT	495
Db	421	AAACAAGTCTTAATCGCATCTGTTTAAACGCTGTTTAAAGCGGCGGCTCTCATCTTAT	480
Qy	496	CAATTAGCTGTCATTTGTCGAATCGCGTTCGCAAAACCATTTCTAGGCGGTGTTTATCG	555
Db	481	CGTTGGCTGTAGCTGTTGAGATGCAATGTTAAGCAAACTCTAGGCGATGTTTATCG	540
Qy	556	---CTTGGCGCAATCAGGTGCTTACAAGAACTCTGAGCTTTTAAACAGTCTCTGTGGC	612
Db	541	TCTGTGGTGGTAAAGTAGCGCTTAAAAAACTCTAGACATTTTAGTGGCCCTATTGGT	600
Qy	613	TGATCATTACAGCGGTATGACAGCGATTGATATTCAGGCGCGCTTATAGGTAACC	672
Db	601	TGGGTCAATTACGCGCGCTTGTAGTACATCAATCTTGTGGCGCGCTTATAGGTAACC	660
Qy	673	ATACGCGCATGCTTGTGTTGCGCACTTTACGCTTAAATAACACAGCAAG	721
Db	661	GTGCTGCTGCTGTTTGTAGTTGCCACCTTACGCAAAATTAAGACAG	709

RESULT 12

US-09-882-227-413
; Sequence 413, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 413
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34) ... (498)
US-09-882-227-413

Query Match 35.8%; Score 271.4; DB 3; Length 547;
Best Local Similarity 71.6%; Pred. No. 3.6e-57;
Matches 389; Conservative 0; Mismatches 141; Indels 13; Gaps 2;

Qy	84	TTTTGTTAAAGACGGCGAAAGACACAAATGAAACCTGACAGCTCCATAGATACAA	143
Db	9	TATGATGAGATGGCAGCACTAAGATGATGAGATTTGACAAATTCACAGATATAA	68
Qy	144	AAGCATGCGCATGATTAGCTTAAATACGCAGAAAGATTCGCTGAAGAGTTTGAATCTA	203
Db	69	AAGATATGCGCATGATTAGCTTAAATACGCAGAAAGATTCGCTGAAGAGTTTGAATCTA	128
Qy	204	TGGGAGCAATAGTTTTCGAGTTTTCATTAACGCGAGGAGTCTTATACAAAGAGATTTT	263

Db 129 TGGGGGCAATAGTTTGGCAATTTTTTTAGATGAAGGGGTCTTATACAAAGAGATTTT 188
Qy 264 ATGCATGTGCGGATAAATTAAGGTCAATTAACAAGAAACTGAAACGACCTTTAAT 323
Db 189 GTGCGATCGTGGCATCATTTAAGGTTAATTAACAAGATCTGCAACCTCTTTGAT 248
Qy 324 TGAACAAACATGCTTTCTAAATCTTAGAAGAAAGTTTGAAGAAATGATGATGAAGA 383
Db 249 TGAGCAAAACATGCTTTCTAACTCTGAAGATAGTTTAGAAAAAATGAGTAGGAGAGA 308
Qy 384 AGTGAAGAAATGCGGATGATATCATTAATAAACACAGGACAA-----TTTAAA 434
Db 309 GATTAAAGAACTTTGCAATGAATGGGCATGCAAAATATGATAAAGTGAATGGTGAAGA 368
Qy 435 CAGACAAGCCCTTAAGCGGGGACCTTTAAACGCTGTTTAAATGGGGGTTTTAAATCTTA 494
Db 369 CAACAAGTCTTAATCGCATCTACTTTAAGCTGTTTAAAGCGGTGGCTCTCATCTTA 428
Qy 495 TCAATTAGCTGTCAATGTTGCGAATGCGGTGCGAAAAACCAATCTAGGCGGTGTTTATC 554
Db 429 TGGCTTGGCTGTATCTGTTGAGATGCAATGTAAGACAACTCTAGGCGATGTTAT--- 485
Qy 555 GCTTGGGGGAATCAGGTGCTTCAAGAACTCTGAGCTTTTAAACAGGTCTCTGTTGGCTG 614
Db 486 -GTGTGGGTAAGTAGCACTTAAAAAAACTTTGGGCGTTTTTGGCTGGCCCTATTGGTTG 544
Qy 615 GAT 617
Db 545 GGT 547

RESULT 13

US-09-732-091-36
; Sequence 36, Application US/09732091
; Patent No. US20020107368A1

; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 36
; LENGTH: 153
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-09-732-091-36

Query Match 20.2%; Score 153; DB 3; Length 153;
Best Local Similarity 100.0%; Pred. No. 7.3e-28;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 601 GGTCTGTGTGGCTGGATCAATTAAGGCGTATGACAGCGATGATGATGACGGGCGCGCT 660
Db 1 GGTCTGTGTGGCTGGATCAATTAAGGCGTATGACAGCGATGATGATGACGGGCGCGCT 60
Qy 661 TATAGGGTAACCATACCGGCATGATGTTGGTTCACCTTTACGCTTAAACACACAGCAA 720
Db 61 TATAGGGTAACCATACCGGCATGATGTTGGTTCACCTTTACGCTTAAACACACAGCAA 120
Qy 721 GCCAATGAGATAAGAACTCGTTGCAAAATAGAA 753
Db 121 GCCAATGAGATAAGAACTCGTTGCAAAATAGAA 153

RESULT 14

US-10-433-970-36
; Sequence 36, Application US/10433970
; Publication No. US20040138415A1

; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 36
; LENGTH: 153
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-10-433-970-36

Query Match 20.2%; Score 153; DB 8; Length 153;
Best Local Similarity 100.0%; Pred. No. 7.3e-28;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 601 GGTCTGTGTGGCTGGATCAATTAAGGCGTATGACAGCGATGATGATGACGGGCGCGCT 660
Db 1 GGTCTGTGTGGCTGGATCAATTAAGGCGTATGACAGCGATGATGATGACGGGCGCGCT 60
Qy 661 TATAGGGTAACCATACCGGCATGATGTTGGTTCACCTTTACGCTTAAACACACAGCAA 720
Db 61 TATAGGGTAACCATACCGGCATGATGTTGGTTCACCTTTACGCTTAAACACACAGCAA 120
Qy 721 GCCAATGAGATAAGAACTCGTTGCAAAATAGAA 753
Db 121 GCCAATGAGATAAGAACTCGTTGCAAAATAGAA 153

RESULT 15

US-09-732-091-33
; Sequence 33, Application US/09732091
; Patent No. US20020107368A1

; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 33
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-09-732-091-33

Query Match 15.0%; Score 114; DB 3; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.5e-18;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 GATTACGCTTAATACGACAGAAAGATCGCTGAAGAGTTGCAATCTATGCGAGCAATAGT 216
Db 1 GATTACGCTTAATACGACAGAAAGATCGCTGAAGAGTTGCAATCTATGCGAGCAATAGT 60
Qy 217 TTTGCGAGTTTCATTAAGGCGAAGAGTCTTATACAAAGAGATTTTATGCGAT 270
Db 61 TTTGCGAGTTTCATTAAGGCGAAGAGTCTTATACAAAGAGATTTTATGCGAT 114

Search completed: August 3, 2006, 10:42:35
Job time : 1475 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2006, 20:14:32 ; Search time 164 Seconds
(without alignments)
7330.066 Million cell updates/sec

Title: US-09-732-091-3

Perfect score: 759

Sequence: 1 atggcacaataatgatag.....cggtgcaatagaatccatt 759

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2193277 seqs, 791917567 residues

Total number of hits satisfying chosen parameters: 4386554

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*

- 1: /EMC Celler SID33/ptodata/1/pubpna/US09 NEW PUB.seq:*
- 2: /EMC Celler SID33/ptodata/1/pubpna/US06 NEW PUB.seq:*
- 3: /EMC Celler SID33/ptodata/1/pubpna/US07 NEW PUB.seq:*
- 4: /EMC Celler SID33/ptodata/1/pubpna/US08 NEW PUB.seq:*
- 5: /EMC Celler SID33/ptodata/1/pubpna/PCT_NEW PUB.seq:*
- 6: /EMC Celler SID33/ptodata/1/pubpna/US10 NEW PUB.seq:*
- 7: /EMC Celler SID33/ptodata/1/pubpna/US11 NEW PUB.seq:*
- 8: /EMC Celler SID33/ptodata/1/pubpna/US11 NEW PUB.seq:*
- 9: /EMC Celler SID33/ptodata/1/pubpna/US60_NEW PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.4	6.0	1738	6	US-10-449-902-5401
2	42.6	5.6	4009	7	US-11-218-305-3660
3	41	5.4	1110	8	US-11-266-748A-52434
4	40.4	5.3	1173	7	US-11-348-413-7111
5	40	5.3	780	7	US-11-359-953-20
6	39.4	5.2	1826	8	US-11-216-545-774
7	39.4	5.2	532643	6	US-10-540-898-308
8	39.2	5.2	1000	8	US-11-266-748A-115072
9	39.2	5.2	1000	8	US-11-266-748A-157236
10	39.2	5.2	1000	8	US-11-266-748A-283572
11	39.2	5.2	1000	8	US-11-266-748A-335001
12	39.2	5.2	1000	8	US-11-266-748A-393321
13	39.2	5.2	1000	8	US-11-266-748A-464367
14	39.2	5.2	1815	8	US-11-266-748A-70584
15	39.2	5.2	1815	8	US-11-266-748A-106086
16	39.2	5.2	1815	8	US-11-266-748A-123395
17	39.2	5.2	2522	7	US-11-218-305-24510
18	39.2	5.2	6521	6	US-10-517-441-310
19	39.2	5.2	6521	6	US-10-517-441-584
20	38.8	5.1	11913	6	US-10-517-441-536
21	38.6	5.1	3532	8	US-11-266-748A-28445
22	38.6	5.1	13286	6	US-10-517-441-472
23	38.4	5.1	151052	6	US-10-540-898-503
24	38	5.0	5087	6	US-10-517-441-320

C 25	38	5.0	5087	6	US-10-517-441-380	Sequence 380, App
C 26	38	5.0	5087	6	US-10-517-441-594	Sequence 594, App
C 27	38	5.0	5087	6	US-10-517-441-654	Sequence 654, App
28	37.6	5.0	168	7	US-11-348-413-12181	Sequence 12181, A
C 29	37.6	5.0	2185	6	US-10-449-902-13623	Sequence 13623, A
C 30	37.6	5.0	5084	8	US-11-105-233-92	Sequence 92, App1
C 31	37.6	5.0	14147	6	US-10-517-441-265	Sequence 265, App
32	37.4	4.9	3434	8	US-11-266-748A-27358	Sequence 27358, A
C 33	37.4	4.9	5493	6	US-10-517-441-691	Sequence 691, App
C 34	37.2	4.9	348	8	US-11-217-529-191030	Sequence 191030, A
C 35	37.2	4.9	4930	6	US-10-517-441-354	Sequence 354, App
C 36	37.2	4.9	4930	6	US-10-517-441-628	Sequence 628, App
C 37	37.2	4.9	6001	6	US-10-517-441-511	Sequence 511, App
C 38	37.2	4.9	6001	6	US-10-517-441-785	Sequence 785, App
C 39	37.2	4.9	7833	6	US-10-517-441-350	Sequence 350, App
C 40	37.2	4.9	7833	6	US-10-517-441-624	Sequence 624, App
C 41	37.2	4.9	8666	6	US-10-517-441-284	Sequence 284, App
C 42	37.2	4.9	8666	6	US-10-517-441-558	Sequence 558, App
43	37	4.9	1698	6	US-10-471-571A-919	Sequence 919, App1
44	37	4.9	2126	1	US-09-949-925-49	Sequence 49, App
C 45	37	4.9	13286	6	US-10-517-441-746	Sequence 746, App

ALIGNMENTS

RESULT 1

US-10-449-902-5401

; Sequence 5401, Application US/10449902

; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.

; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF

; FILE REFERENCE: MOA-A02051-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5401

; LENGTH: 1738

; TYPE: DNA

; ORGANISM: Oryza sativa

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: AK062580

; DATABASE ENTRY DATE: 2001-12-06

US-10-449-902-5401

Query Match

Best Local Similarity 6.0%; Score 45.4; DB 6; Length 1738;

Matches 85; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 134 TAGATACAAAAGGCATGCGATGATTACGCTAAATACGCAGAAAGATCGCTGAAGT 193

1258 TTGAACAACATCCGATATGAATGAATGCTGTAATATGATGCGCAATTGGATGAATTT 1317

QY 194 TCGAATATCTGGGAGCAATAGTTTGGAGTTTCAATTAAGCGAAGAGTCTTATACA 253

1318 TGTAACTACTAGAGAAAGAAATACCTTTGGAGTTTCCAATTAAGTGTCAAGTTTATTGTTCA 1377

QY 254 AAGAGATTTTATGCGATGTCGCGATAATT 284

DB 1378 TAGGCACATATTATTGTTTTCACAAATTTGTT 1408

RESULT 2

US-11-218-305-3660

```
Sequence 3660, Application US/11218305
Publication No. US20060141495A1
GENERAL INFORMATION:
APPLICANT: MONSANTO TECHNOLOGY, LLC
APPLICANT: McLaird, Paul L.
APPLICANT: Tao, Nengbing
APPLICANT: Wu, Kunsheng
TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
TITLE OF INVENTION: Corn.
FILE REFERENCE: 38-21 (53660)B
CURRENT APPLICATION NUMBER: US/11/218,305
CURRENT FILING DATE: 2005-09-01
PRIOR APPLICATION NUMBER: US 60/606,880
PRIOR FILING DATE: 2004-09-01
NUMBER OF SEQ ID NOS: 25043
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3660
LENGTH: 4009
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (22)..(22)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (111)..(111)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (178)..(178)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (235)..(235)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (285)..(285)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (290)..(290)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (308)..(308)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (310)..(313)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (341)..(341)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (362)..(364)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (408)..(408)
OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-3660
Query Match 5.6%; Score 42.6; DB 7; Length 4009;
Best Local Similarity 44.1%; Pred. No. 0.31;
Matches 150; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
QY 103 AAAGACACATGAAACCTCACAGCTCCATAGATACAAAGGCATGGCGATGATTAC 162
Db 1 AAAAAAAAAAAGAGGACGCGNAAAAAAAAAAAAAAAAAAAAAAAAAAGAGAGAGAAAAA 60
163 GCTAAATACGCAGAAAGAAATCGCTGAAGAGTTGCAATACTATGGGAGCAATAGTTTGGC 222
61 AATAAAAAAAAAACACAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 120
223 AGTTTTCATTAAAGGCGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAA 282
121 AAGGGGAAACAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 180
283 TTAAGGTCAATTACAAAGAAACTGAAACGACTTTAAATTGAACAAAAACATGCTTTCT 342
181 AAAAAAAAAAATAATAAATTTAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 240
343 AAAATCTTAGAAGAAAGTTTGGAGAAATGGATGATGAAGAGTGAAGAAAAATGTGCGAT 402
241 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 300
403 GAATTATCCATAAAAAACACGCAAAATTTAAACACACAAG 442
301 AAAAAAAAAANNNNAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAG 340
RESULT 3
US-11-266-748A-52434
Sequence 52434, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
METHODS OF USING THE SAME
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 52434
LENGTH: 1110
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (761)..(763)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (786)..(786)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (790)..(791)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (802)..(807)
OTHER INFORMATION: n is a, c, g, or t
```

```
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (811)..(811)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (814)..(815)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (827)..(827)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (837)..(842)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (847)..(851)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (853)..(853)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (855)..(855)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (860)..(860)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (862)..(866)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (873)..(873)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (875)..(875)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (877)..(881)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (888)..(890)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (894)..(898)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (902)..(903)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (911)..(912)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (916)..(924)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (933)..(937)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
```

```
/ NAME/KEY: misc feature
/ LOCATION: (946)..(947)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (969)..(971)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (973)..(976)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1005)..(1005)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1029)..(1029)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1041)..(1041)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1051)..(1052)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1086)..(1086)
/ OTHER INFORMATION: n is a, c, g, or t
/ US-11-266-748A-52434
```

```
Query Match 5.4%; Score 41; DB 8; Length 1110;
Best Local Similarity 37.6%; Pred. No. 0.51;
Matches 74; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
```

```
QY 282 ATTAAGGTCATTAACAACAAGAACTGAAACGACTTTTAAATTGAAACAACATGCTTTC 341
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
748 ATCAAAAAAANNNNAAAAAANNNNAAAAAANNNNAAAAAANNNNAAAAAANNNNN 807
QY 342 TAAATCTTAGAAAGAAAGTTTGGAGAAATGGATGATGAAGAGTGAAGAAATGTCGA 401
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
808 AAANAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANNNNA 867
QY 402 TGAATTATCCATAAAAAACACGACCAATTTAAACAGACAGCCTTAAAGCCGCGACTTT 461
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
868 AAAAAANANNNNNAAAAAANNNNAAAAAANNNNAAAAAANNNNAAAAAANNNNNAAA 927
QY 462 AACGCTGTTTAAATGG 478
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
928 AAAAAANNNNAAAAAAGG 944
```

```
RESULT 4
US-11-348-413-7111
; Sequence 7111, Application US/11348413
; Publication No. US20060160121A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
```

```

; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 7111
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1173)
; OTHER INFORMATION: WAN01ULX2; ORF: 00000011100000; Cluster contains WAN01PJB8
; OTHER INFORMATION: SAG1752::conserved hypothetical protein TIGR00275:2603:NC
; OTHER INFORMATION: _004116.1
US-11-348-413-7111

```

```

Query Match      5.3%; Score 40.4; DB 7; Length 1173;
Best Local Similarity 45.1%; Pred. No. 0.75;
Matches 183; Conservative 2; Mismatches 218; Indels 3; Gaps 1;

QY 218 TTTCGAGTTTCATTAAAGCGGAGGAGTCTTTATACAAAGAGATTTTATGCGATGTGTGCG 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 TTTCYAGCTTTGTCAAGGGGAGAACTATATACCTAGATGTTTTCGCCCAATATCTCTG 784

QY 278 ATAAATTAAGGTCAATATCAACAAGAAACTGAAAGCACTTTAATTTGAACAAAAACATGC 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 TTAAGAATTTGGAAYACACTTC--CAAAATGAAAGAGAAAAATCTTTGAAAAACGCTC 841

QY 338 TTTCCTAAATCTTAGAAGAGTTTGGAGAAATGATGATGAAGAGTGAAGAAATGT 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 TTAATAATCTTCTTCCTGAGCGTTTAGCTGAATTTTACGCTGAAGACTTACCMGAAAAAG 901

QY 398 GCATGAATTTATCCATAAAAAACACGACAAATTTAAACAGACAGCGCTTAAAGCGCGCGA 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 TAAACACAGTACTGTAAAGACTTAGAAYGTTAATTCAAAACTTAAAAAATCTCTTA 961

QY 458 CTTTAACGCTGTTTAAATGGGGGTTTAAATCTTATCAATTAAGCTGCTGATGTTGCGA 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 962 TTTTAGTAACCTGGAAGATCTCTCTAGCAAAATCTTTTGAACAAAAAGGAGGAGTTGATC 1021

QY 518 ATCCGCTGCGAAAAACCTTCTAGGCGTGTGTTATGCTTGGCGGCAATCAGTGTCTTA 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1022 TTAAGAATAATAAACCCCAACACTTGAAGTAAAAAGGAGTGTGTTTACATTTTCTGCT 1081

QY 578 CAAGAACTCTGAGCTTTTAAACAGGTCTCTGCTGATGATCAATAC 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1082 GTGAAGTCTAGACATCAATGCTCATACGGAGGGTTTAACTTAC 1127

```

```

RESULT 5
US-11-359-953-20
; Sequence 20, Application US/11359953
; Publication No. US20060141577A1
; GENERAL INFORMATION:
; APPLICANT: ChromaGenics B.V.
; APPLICANT: Otte, Arie P.
; APPLICANT: Kwaks, Theodorus H.J.
; APPLICANT: Sewalt, Richard G.A.B.
; APPLICANT: van Blootland, Rik
; TITLE OF INVENTION: Selection of host cells expressing protein at high levels
; FILE REFERENCE: 0117 A US P00 CIP
; CURRENT APPLICATION NUMBER: US/11/359,953
; CURRENT FILING DATE: 2006-02-21
; PRIOR APPLICATION NUMBER: US 11/269,525
; PRIOR FILING DATE: 2005-11-07
; PRIOR APPLICATION NUMBER: US 60/626,301
; PRIOR FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: US 60/696,610
; PRIOR FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: EP 04105593.0
; PRIOR FILING DATE: 2004-11-08
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 780
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: sequence of STAR20
US-11-359-953-20

Query Match      5.3%; Score 40; DB 7; Length 780;
Best Local Similarity 53.1%; Pred. No. 0.82;
Matches 85; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 246 CTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATAAAGGTCAATTACCAACAGAA 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 CTCTGACATAGATATTTAAATTTCTTAGTGGTTTGTGTATATTTCTTATTAATA 538

QY 306 AACTGAAACGACTTTAATTTGAACAAAAACATGCTTTCTAAAATCTTAGAAAGAGTTTGA 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 AAGCATTTATTTGTTTACAGAAAAAAGATATATATCTTAAATCTTAAAAATAAATAACCA 598

QY 366 AGAAATGATGATGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGTGAAGAAGTGAAGA 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 TTAAGAGAAAAACAGGAGTTTAACTAATAAGGGAACAA 638

RESULT 6
US-11-216-545-774
; Sequence 774, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Soybeans
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 774
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15)..(16)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (33)..(39)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (42)..(43)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45)..(46)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(53)
; OTHER INFORMATION: n is a, c, g, or t

```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (58)..(61)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (73)..(73)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (75)..(80)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (97)..(97)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100)..(103)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (106)..(109)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (112)..(113)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (118)..(121)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (125)..(125)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (133)..(133)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (136)..(141)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148)..(152)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (156)..(158)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (171)..(172)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (175)..(175)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (190)..(192)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (196)..(198)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (214)..(218)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (224)..(224)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (226)..(228)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (230)..(232)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (247)..(248)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (252)..(252)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (254)..(255)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (263)..(264)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (269)..(273)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (275)..(277)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (283)..(285)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (288)..(288)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (291)..(293)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (297)..(297)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (300)..(301)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (308)..(310)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (313)..(317)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (322)..(324)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (331)..(331)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
```


;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105507.0
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105485.9
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: EP 04105484.2
;; PRIOR FILING DATE: 2004-11-03
;; PRIOR APPLICATION NUMBER: US 60/662,276
;; PRIOR FILING DATE: 2005-03-14
;; PRIOR APPLICATION NUMBER: US 60/700,293
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 157236
;; LENGTH: 1000
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
US-11-266-748A-157236

Query Match 5.2%; Score 39.2; DB 8; Length 1000;
Best Local Similarity 55.0%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 8 ACAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGAATTTATTTGGATT 67
DB 766 AAAAAGAAGAAGAAATGGCAAGAGAAAGTTTTTCAAAATTTCTTTCTTTTAAATT 825
QY 68 TGTGAGGTCTGTTGTTTGTAAAGACGGCGGCAAAAGACACACATGAAAACCTGACCA 127
DB 826 TAGATTGAGTTTCATTTATTTGAAACAGACTGGGCCAATGTCCACAAAGAAATTCCTGGTCA 885
QY 128 GCTCCATAGATACAAAGG 147
DB 886 GCACCACCGATGTCCAAAGG 905

RESULT 10

US-11-266-748A-283572/c
; Sequence 283572, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 283572
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-283572

Query Match 5.2%; Score 39.2; DB 8; Length 1000;
Best Local Similarity 55.0%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 8 ACAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGAATTTATTTGGATT 67
DB 235 AAAAAGAAGAAGAAATGGCAAGAGAAAGTTTTTCAAAATTTCTTTCTTTTAAATT 176
QY 68 TGTGAGGTCTGTTGTTTGTAAAGACGGCGGCAAAAGACACACATGAAAACCTGACCA 127
DB 175 TAGATTGAGTTTCATTTATTTGAAACAGACTGGGCCAATGTCCACAAAGAAATTCCTGGTCA 116
QY 128 GCTCCATAGATACAAAGG 147
DB 115 GCACCACCGATGTCCAAAGG 96

RESULT 11

US-11-266-748A-335001
; Sequence 335001, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 335001
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-335001

Query Match 5.2%; Score 39.2; DB 8; Length 1000;
Best Local Similarity 55.0%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 8 ACAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGAATTTATTTGGATT 67
DB 766 AAAAAGAAGAAGAAATGGCAAGAGAAAGTTTTTCAAAATTTCTTTCTTTTAAATT 825
QY 68 TGTGAGGTCTGTTGTTTGTAAAGACGGCGGCAAAAGACACACATGAAAACCTGACCA 127
DB 826 TAGATTGAGTTTCATTTATTTGAAACAGACTGGGCCAATGTCCACAAAGAAATTCCTGGTCA 885
QY 128 GCTCCATAGATACAAAGG 147
DB 886 GCACCACCGATGTCCAAAGG 905

RESULT 12
US-11-266-748A-393321/c

```
; Sequence 393321, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 393321
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-11-266-748A-393321

Query Match          5.2%; Score 39.2; DB 8; Length 1000;
Best Local Similarity 55.0%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy      8  ACAATATGATAGAGACTTGGAAATTTTAAAGCAATTGGAATCTAGTGTATTGGATT 67
Db      235 AAAAAGAGAGAGAAAGAAATGCAAGAGAGAAAGTTTTCAAATTTCTTTCTTTTAAAT 176
Qy      68  TGTTTGAGTCTTGTCTTTTGTAAAGACGGCGAAAGACACAAATGAAAACTGACCA 127
Db      175 TAGATTGAGTTCAATTATTGAAACAGACTGGGCAATGTCCACAAAGAAATTCCTGTGTC 116
Qy      128  GCTCCATAGATACAAAAGG 147
Db      115 GCACCACCGATGTCCAAAGG 96

RESULT 13
; Sequence 464367, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 464367
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-11-266-748A-464367

Query Match          5.2%; Score 39.2; DB 8; Length 1000;
Best Local Similarity 55.0%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy      8  ACAATATGATAGAGACTTGGAAATTTTAAAGCAATTGGAATCTAGTGTATTGGATT 67
Db      235 AAAAAGAGAGAGAAAGAAATGCAAGAGAGAAAGTTTTCAAATTTCTTTCTTTTAAAT 176
Qy      68  TGTTTGAGTCTTGTCTTTTGTAAAGACGGCGAAAGACACAAATGAAAACTGACCA 127
Db      175 TAGATTGAGTTCAATTATTGAAACAGACTGGGCAATGTCCACAAAGAAATTCCTGTGTC 116
Qy      128  GCTCCATAGATACAAAAGG 147
Db      115 GCACCACCGATGTCCAAAGG 96

RESULT 14
; Sequence 70584, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 70584
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: (1190)..(1198)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
```

```
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 464367
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-11-266-748A-464367

Query Match          5.2%; Score 39.2; DB 8; Length 1000;
Best Local Similarity 55.0%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy      8  ACAATATGATAGAGACTTGGAAATTTTAAAGCAATTGGAATCTAGTGTATTGGATT 67
Db      766 AAAAAGAGAGAGAAAGAAATGCAAGAGAGAAAGTTTTCAAATTTCTTTCTTTTAAAT 825
Qy      68  TGTTTGAGTCTTGTCTTTTGTAAAGACGGCGAAAGACACAAATGAAAACTGACCA 127
Db      826 TAGATTGAGTTCAATTATTGAAACAGACTGGGCAATGTCCACAAAGAAATTCCTGTGTC 885
Qy      128  GCTCCATAGATACAAAAGG 147
Db      886 GCACCACCGATGTCCAAAGG 905

RESULT 14
; Sequence 70584, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 70584
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: (1190)..(1198)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
```

:	NAME/KEY:	misc feature			
:	LOCATION:	(1248)-(1271)			
:	OTHER INFORMATION:	n is a, c, g, or t			
US-11-266-748A-70584					
 Query Match 5.2%; Score 39.2; DB 8; Length 1815;					
Best Local Similarity 55.0%; Pred. No. 1.8;					
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;					
Qy	8	ACAAATATGATAGACTTCGCAATTTTAAAGCAAATTGGAACTCTAGTGATTATTGGATT	67		
Db	1049	AAAAAGAGAGAAGAAATGC			
Qy	68	TGTTTGAGGTGCTTGT			
Db	989	TAGATTGAGTTCATTTATTTGAACACAGACTGGGCCAATGTCCACAAGAATTCCTGGTCA	930		
 Search completed: August 4, 2006, 20:21:10					
Job time : 166 secs					

```

RESULT 15
US-11-266-748A-106086/c
; Sequence 106086, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 106086
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1190)..(1198)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1248)..(1271)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-106086

Query Match 5.2%; Score 39.2; DB 8; Length 1815;
Best Local Similarity 55.0%; Pred. No. 1.8;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

8 ACAAAATATGATGAGACTTGGAAATTTTAAAGCAATCTAGTGATTTATTGGATT 67

```

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**

Because they contain data that is confidential, the results of Pending database searches should not be left in the case .

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2006, 09:12:56 ; Search time 5922 Seconds
(without alignments)
7492.404 Million cell updates/sec

Title: US-09-732-091-3

Perfect score: 759

Sequence: 1 atggcatacaaatgatag.....cgttgcaaatagaatccatt 759

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 86534536 seqs, 29229259966 residues

Total number of hits satisfying chosen parameters: 173069072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pna/PCTUSA COMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pna/PCTUSB COMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pna/PCTUSC COMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pna/US075 COMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pna/US076 COMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pna/US077 COMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pna/US078 COMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pna/US079 COMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pna/US080 COMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pna/US081 COMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pna/US082 COMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pna/US083 COMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pna/US084 COMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pna/US085 COMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pna/US086 COMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pna/US087 COMB.seq.*
- 17: /EMC_Celerra_SIDS3/ptodata/2/pna/US088 COMB.seq.*
- 18: /EMC_Celerra_SIDS3/ptodata/2/pna/US089 COMB.seq.*
- 19: /EMC_Celerra_SIDS3/ptodata/2/pna/US090 COMB.seq.*
- 20: /EMC_Celerra_SIDS3/ptodata/2/pna/US091 COMB.seq.*
- 21: /EMC_Celerra_SIDS3/ptodata/2/pna/US092 COMB.seq.*
- 22: /EMC_Celerra_SIDS3/ptodata/2/pna/US093 COMB.seq.*
- 23: /EMC_Celerra_SIDS3/ptodata/2/pna/US094 COMB.seq.*
- 24: /EMC_Celerra_SIDS3/ptodata/2/pna/US095A COMB.seq.*
- 25: /EMC_Celerra_SIDS3/ptodata/2/pna/US095B COMB.seq.*
- 26: /EMC_Celerra_SIDS3/ptodata/2/pna/US095C COMB.seq.*
- 27: /EMC_Celerra_SIDS3/ptodata/2/pna/US096A COMB.seq.*
- 28: /EMC_Celerra_SIDS3/ptodata/2/pna/US096B COMB.seq.*
- 29: /EMC_Celerra_SIDS3/ptodata/2/pna/US096C COMB.seq.*
- 30: /EMC_Celerra_SIDS3/ptodata/2/pna/US097A COMB.seq.*
- 31: /EMC_Celerra_SIDS3/ptodata/2/pna/US097B COMB.seq.*
- 32: /EMC_Celerra_SIDS3/ptodata/2/pna/US098A COMB.seq.*
- 33: /EMC_Celerra_SIDS3/ptodata/2/pna/US098B COMB.seq.*
- 34: /EMC_Celerra_SIDS3/ptodata/2/pna/US099A COMB.seq.*
- 35: /EMC_Celerra_SIDS3/ptodata/2/pna/US099B COMB.seq.*
- 36: /EMC_Celerra_SIDS3/ptodata/2/pna/US099C COMB.seq.*
- 37: /EMC_Celerra_SIDS3/ptodata/2/pna/US099D COMB.seq.*
- 38: /EMC_Celerra_SIDS3/ptodata/2/pna/US099E COMB.seq.*
- 39: /EMC_Celerra_SIDS3/ptodata/2/pna/US100A COMB.seq.*
- 40: /EMC_Celerra_SIDS3/ptodata/2/pna/US100B COMB.seq.*
- 41: /EMC_Celerra_SIDS3/ptodata/2/pna/US101 COMB.seq.*
- 42: /EMC_Celerra_SIDS3/ptodata/2/pna/US102A COMB.seq.*
- 43: /EMC_Celerra_SIDS3/ptodata/2/pna/US102B COMB.seq.*

- 44: /EMC_Celerra_SIDS3/ptodata/2/pna/US103A COMB.seq.*
- 45: /EMC_Celerra_SIDS3/ptodata/2/pna/US103B COMB.seq.*
- 46: /EMC_Celerra_SIDS3/ptodata/2/pna/US103C COMB.seq.*
- 47: /EMC_Celerra_SIDS3/ptodata/2/pna/US103D COMB.seq.*
- 48: /EMC_Celerra_SIDS3/ptodata/2/pna/US103E COMB.seq.*
- 49: /EMC_Celerra_SIDS3/ptodata/2/pna/US103F COMB.seq.*
- 50: /EMC_Celerra_SIDS3/ptodata/2/pna/US104 COMB.seq.*
- 51: /EMC_Celerra_SIDS3/ptodata/2/pna/US105 COMB.seq.*
- 52: /EMC_Celerra_SIDS3/ptodata/2/pna/US106A COMB.seq.*
- 53: /EMC_Celerra_SIDS3/ptodata/2/pna/US106B COMB.seq.*
- 54: /EMC_Celerra_SIDS3/ptodata/2/pna/US107A COMB.seq.*
- 55: /EMC_Celerra_SIDS3/ptodata/2/pna/US107B COMB.seq.*
- 56: /EMC_Celerra_SIDS3/ptodata/2/pna/US107C COMB.seq.*
- 57: /EMC_Celerra_SIDS3/ptodata/2/pna/US107D COMB.seq.*
- 58: /EMC_Celerra_SIDS3/ptodata/2/pna/US107E COMB.seq.*
- 59: /EMC_Celerra_SIDS3/ptodata/2/pna/US107F COMB.seq.*
- 60: /EMC_Celerra_SIDS3/ptodata/2/pna/US107G COMB.seq.*
- 61: /EMC_Celerra_SIDS3/ptodata/2/pna/US108 COMB.seq.*
- 62: /EMC_Celerra_SIDS3/ptodata/2/pna/US109A COMB.seq.*
- 63: /EMC_Celerra_SIDS3/ptodata/2/pna/US109B COMB.seq.*
- 64: /EMC_Celerra_SIDS3/ptodata/2/pna/US109C COMB.seq.*
- 65: /EMC_Celerra_SIDS3/ptodata/2/pna/US110A COMB.seq.*
- 66: /EMC_Celerra_SIDS3/ptodata/2/pna/US110B COMB.seq.*
- 67: /EMC_Celerra_SIDS3/ptodata/2/pna/US110C COMB.seq.*
- 68: /EMC_Celerra_SIDS3/ptodata/2/pna/US110D COMB.seq.*
- 69: /EMC_Celerra_SIDS3/ptodata/2/pna/US111A COMB.seq.*
- 70: /EMC_Celerra_SIDS3/ptodata/2/pna/US111B COMB.seq.*
- 71: /EMC_Celerra_SIDS3/ptodata/2/pna/US112A COMB.seq.*
- 72: /EMC_Celerra_SIDS3/ptodata/2/pna/US112B COMB.seq.*
- 73: /EMC_Celerra_SIDS3/ptodata/2/pna/US113A COMB.seq.*
- 74: /EMC_Celerra_SIDS3/ptodata/2/pna/US113B COMB.seq.*
- 75: /EMC_Celerra_SIDS3/ptodata/2/pna/US114 COMB.seq.*
- 76: /EMC_Celerra_SIDS3/ptodata/2/pna/US117 COMB.seq.*
- 77: /EMC_Celerra_SIDS3/ptodata/2/pna/US600 COMB.seq.*
- 78: /EMC_Celerra_SIDS3/ptodata/2/pna/US601 COMB.seq.*
- 79: /EMC_Celerra_SIDS3/ptodata/2/pna/US602A COMB.seq.*
- 80: /EMC_Celerra_SIDS3/ptodata/2/pna/US602B COMB.seq.*
- 81: /EMC_Celerra_SIDS3/ptodata/2/pna/US603 COMB.seq.*
- 82: /EMC_Celerra_SIDS3/ptodata/2/pna/US604A COMB.seq.*
- 83: /EMC_Celerra_SIDS3/ptodata/2/pna/US604B COMB.seq.*
- 84: /EMC_Celerra_SIDS3/ptodata/2/pna/US605 COMB.seq.*
- 85: /EMC_Celerra_SIDS3/ptodata/2/pna/US606 COMB.seq.*
- 86: /EMC_Celerra_SIDS3/ptodata/2/pna/US607 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	759	100.0	759	1	PCT-US01-48392-3	Sequence 3, Appli
2	759	100.0	759	30	US-09-732-091-3	Sequence 3, Appli
3	759	100.0	759	50	US-10-433-970-3	Sequence 3, Appli
4	759	100.0	795	1	PCT-US01-48392-43	Sequence 43, Appl
5	759	100.0	795	30	US-09-732-091-43	Sequence 43, Appl
6	759	100.0	795	50	US-10-433-970-43	Sequence 43, Appl
7	727	95.8	762	3	PCT-US96-18542-91	Sequence 91, Appl
8	727	95.8	762	14	US-08-561-469A-971	Sequence 971, App
9	727	95.8	762	16	US-08-761-184-48	Sequence 48, Appl
10	727	95.8	762	16	US-08-761-184-649	Sequence 649, App
11	727	95.8	762	17	US-08-821-931-48	Sequence 48, Appl
12	727	95.8	762	17	US-08-821-931-649	Sequence 649, App
13	727	95.8	762	18	US-08-993-002A-4400	Sequence 4400, App
14	727	95.8	762	18	US-08-993-002A-4401	Sequence 4401, Ap
15	727	95.8	762	49	US-10-335-977-4400	Sequence 4400, Ap
16	727	95.8	762	49	US-10-335-977-4401	Sequence 4401, Ap
17	727	95.8	768	16	US-08-761-184-393	Sequence 393, App
18	727	95.8	768	17	US-08-821-931-393	Sequence 393, App
19	727	95.8	768	18	US-08-993-002A-4402	Sequence 4402, Ap
20	727	95.8	768	49	US-10-335-977-4402	Sequence 4402, Ap

Sequence 312, App
Sequence 47, Appl
Sequence 47, Appl
Sequence 45, Appl
Sequence 45, Appl
Sequence 2937, Ap
Sequence 2936, Ap
Sequence 2936, Ap
Sequence 1147, Ap
Sequence 413, App
Sequence 413, App
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 33, Appl
Sequence 33, Appl
Sequence 35, Appl
Sequence 35, Appl
Sequence 32, Appl
Sequence 34, Appl
Sequence 32, Appl

15 US-08-621-425-312
1 PCT-US01-48392-47
50 US-10-433-970-47
1 PCT-US01-48392-45
50 US-10-433-970-45
18 US-08-993-002A-2937
49 US-10-335-977-2937
18 US-08-993-002A-2936
49 US-10-335-977-2936
3 PCT-US98-06371-1147
18 US-08-902-615A-413
33 US-09-882-227-413
1 PCT-US01-48392-36
30 US-09-732-091-36
153 20.2 153 30 US-09-732-091-36
153 20.2 153 30 US-10-433-970-36
151.6 20.0 279 15 US-08-625-811-582
114 15.0 114 1 PCT-US01-48392-33
114 15.0 114 30 US-09-732-091-33
114 15.0 114 50 US-10-433-970-33
108 14.2 108 1 PCT-US01-48392-35
108 14.2 108 30 US-09-732-091-35
108 14.2 108 50 US-10-433-970-35
90 11.9 90 1 PCT-US01-48392-32
90 11.9 90 1 PCT-US01-48392-34
90 11.9 90 30 US-09-732-091-32

ALIGNMENTS

RESULT 1

PCT-US01-48392-3
; Sequence 3, Application PC/TUS0148392
; GENERAL INFORMATION:
; APPLICANT: AntexBiologicals, Inc.
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; FILE REFERENCE: 7969-091-228
; CURRENT APPLICATION NUMBER: PCT/US01/48392
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Helicobacter sp.
PCT-US01-48392-3

Query Match 100.0%; Score 759; DB 1; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.4e-221;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGATCAAAATATGATAGAGACTTGAATTTTAAAGCAATTCGAATCTAGTATTTA 60
Db 1 ATGCGATCAAAATATGATAGAGACTTGAATTTTAAAGCAATTCGAATCTAGTATTTA 60
Qy 61 TTGGATTTGTTGAGGTGCTTTTTCGTAAGACGGCGGAAAGACACAATGAAAAA 120
Db 61 TTGGATTTGTTGAGGTGCTTTTTCGTAAGACGGCGGAAAGACACAATGAAAAA 120
Qy 121 CTGACCAAGCTCCATAGATAAAGGCGATGCGGATGATTACGCTAAATACGCAAGAGA 180
Db 121 CTGACCAAGCTCCATAGATAAAGGCGATGCGGATGATTACGCTAAATACGCAAGAGA 180
Qy 181 ATCGCTGAAGAGTTGCAATCTATGGAGCAATAGTTTTTGGAGTTTCATTAAGGCGAA 240
Db 181 ATCGCTGAAGAGTTGCAATCTATGGAGCAATAGTTTTTGGAGTTTCATTAAGGCGAA 240
Qy 241 GGAGTCTTATCAAGAGATTTTATGCGATGTGTCGATTAATTAAGGTCAATTAACAC 300
Db 241 GGAGTCTTATCAAGAGATTTTATGCGATGTGTCGATTAATTAAGGTCAATTAACAC 300

Qy 301 AAGAAACTGAACGACTTTTAATTGAACAAACATGCTTTCTAAATCTTAGAAGAAGT 360
Db 301 AAGAAACTGAACGACTTTTAATTGAACAAACATGCTTTCTAAATCTTAGAAGAAGT 360
Qy 361 TTGGAAGAAATGGATGATGAAGAGTGAAGAAATGTGCGATGAATATCCATAAAAAAC 420
Db 361 TTGGAAGAAATGGATGATGAAGAGTGAAGAAATGTGCGATGAATATCCATAAAAAAC 420
Qy 421 ACGGACAAATTTAAACAGACAGCCCTTAAGCGCGGCGACTTTAAACGCTGTTTAAATGGGG 480
Db 421 ACGGACAAATTTAAACAGACAGCCCTTAAGCGCGGCGACTTTAAACGCTGTTTAAATGGGG 480
Qy 481 GGTCTTAAATCTTCAATTAAGTCTCATTTGTTGCGAATCGGTCCCAAAACCATTTCTA 540
Db 481 GGTCTTAAATCTTCAATTAAGTCTCATTTGTTGCGAATCGGTCCCAAAACCATTTCTA 540
Qy 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTCAAGAACTCTGAGCTTTTAAACA 600
Db 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTCAAGAACTCTGAGCTTTTAAACA 600
Qy 601 GGTCTGTTGGCTGGATCATTTACAGCGGTATGGACAGCAATTCATATTCAGGGCGGCT 660
Db 601 GGTCTGTTGGCTGGATCATTTACAGCGGTATGGACAGCAATTCATATTCAGGGCGGCT 660
Qy 661 TATAGGTAACCAATACCGGCGATGCTGTTGTTGCCACTTTAGCCCTTAAACACACAGCAA 720
Db 661 TATAGGTAACCAATACCGGCGATGCTGTTGTTGCCACTTTAGCCCTTAAACACACAGCAA 720
Qy 721 GCCAATGAGATAGAGAGTTCGTTGCAATAGAAATCCATT 759
Db 721 GCCAATGAGATAGAGAGTTCGTTGCAATAGAAATCCATT 759

RESULT 2

US-09-732-091-3
; Sequence 3, Application US/09732091
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-09-732-091-3

Query Match 100.0%; Score 759; DB 30; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.4e-221;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGATCAAAATATGATAGAGACTTGAATTTTAAAGCAATTCGAATCTAGTATTTA 60
Db 1 ATGCGATCAAAATATGATAGAGACTTGAATTTTAAAGCAATTCGAATCTAGTATTTA 60
Qy 61 TTGGATTTGTTGAGGTGCTTTTTCGTAAGACGGCGGAAAGACACAATGAAAAA 120
Db 61 TTGGATTTGTTGAGGTGCTTTTTCGTAAGACGGCGGAAAGACACAATGAAAAA 120
Qy 121 CTGACCAAGCTCCATAGATAAAGGCGATGCGGATGATTACGCTAAATACGCAAGAGA 180
Db 121 CTGACCAAGCTCCATAGATAAAGGCGATGCGGATGATTACGCTAAATACGCAAGAGA 180
Qy 181 ATCGCTGAAGAGTTGCAATCTATGGAGCAATAGTTTTTGGAGTTTCATTAAGGCGAA 240
Db 181 ATCGCTGAAGAGTTGCAATCTATGGAGCAATAGTTTTTGGAGTTTCATTAAGGCGAA 240

241	Qy	GGAGTCTTTATACAAAGAGATTTTATGCGATGTGTCGATAAAATTTAAAGGTCAATTACAACT	300
241	Db	GGAGTCTTTATACAAAGAGATTTTATGCGATGTGTCGATAAAATTTAAAGGTCAATTACAACT	300
301	Qy	AAGAAACTGAAAACGACTTTTAATTCGAAACAAACATGCTTTCTTAAATCTTTAGAAAGAAAGT	360
301	Db	AAGAAACTGAAAACGACTTTTAATTCGAAACAAACATGCTTTCTTAAATCTTTAGAAAGAAAGT	360
361	Qy	TTGGAAGAAATGGATGATGGAAGAGTGAAGAAATGTGCGATCAATATATCCATATAAAAC	420
361	Db	TTGGAAGAAATGGATGATGGAAGAGTGAAGAAATGTGCGATCAATATATCCATATAAAAC	420
421	Qy	ACGGAACAATTTAAACAGACAACAGCCTTTAAGCGCGCGAATTTTAACGCTGTTTAAAAATGGGG	480
421	Db	ACGGAACAATTTAAACAGACAACAGCCTTTAAGCGCGCGAATTTTAACGCTGTTTAAAAATGGGG	480
481	Qy	GGTTTTAAATCTTATCAANTTAGCTGTCATGTTTGCGAATGCGGTGCGAAAAACCATTTCTA	540
481	Db	GGTTTTAAATCTTATCAANTTAGCTGTCATGTTTGCGAATGCGGTGCGAAAAACCATTTCTA	540
541	Qy	GGGCGTGGTTTTATCGCTTTCGGGCGCAATCAGGTGCTTTACAAGAACTCTGAGCTTTTTTAACA	600
541	Db	GGGCGTGGTTTTATCGCTTTCGGGCGCAATCAGGTGCTTTACAAGAACTCTGAGCTTTTTTAACA	600
601	Qy	GGTCTGTTGGCTGGATCAATTTACAGGCGGTATGGACAGCGATTTGATATTGACGGGCGCGCT	660
601	Db	GGTCTGTTGGCTGGATCAATTTACAGGCGGTATGGACAGCGATTTGATATTGACGGGCGCGCT	660
661	Qy	TATAGGCTTAACCATACCGGCATGCAATTCGTGTTGCCACTTTACGCTTAAAAACACAGCAA	720
661	Db	TATAGGCTTAACCATACCGGCATGCAATTCGTGTTGCCACTTTACGCTTAAAAACACAGCAA	720
721	Qy	GCCAAATGGAGATAAGAAGTCGTTGCCAAATAGAATCCATT	759
721	Db	GCCAAATGGAGATAAGAAGTCGTTGCCAAATAGAATCCATT	759

```

RESULT 3
US-10-433-970-3
; Sequence 3, Application US/10433970
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Helicobacter sp.
US-10-433-970-3

```

Query Match	100.0%	Score 759;	DB 50;	Length 759;
Best Local Similarity	100.0%	Pred. No. 1.4e-221;		
Matches 759;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

1	ATGCATACAAATATGATAGAGACTTTGGAATTTTTAAAGCAATTCGAATCTAGTGATTTA	60	Qy
1	ATGCATACAAATATGATAGAGACTTTGGAATTTTTAAAGCAATTCGAATCTAGTGATTTA	60	Db
61	TTGGAATTGTTTGAGTGCTGTTTTTGGTAAAGACGGCGAAAAAAGACACAATGAAAAA	120	Qy
61	TTGGAATTGTTTGAGTGCTGTTTTTGGTAAAGACGGCGAAAAAAGACACAATGAAAAA	120	Db
121	CTGACCCAGCTCCATAGAAATACAAAGGCGATCGCGATGATTACGGCTAAATACGCAGAAAGA	180	Qy

Db	121	CTGACCAGCTCCATAGAAATACAAAAGGATGCGGATGATTACGCTAAATACCGACGAAGA	180
Qy	181	ATCGCTGAAGAGTTGCAATACTATGGGAGCAATAGTTTTCGAGTTTCATTAAAGCGCGAA	240
Db	181	ATCGCTGAAGAGTTGCAATACTATGGGAGCAATAGTTTTCGAGTTTCATTAAAGCGCGAA	240
Qy	241	GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGGATAAATTAAGGTCAATTACAAAC	300
Db	241	GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGGATAAATTAAGGTCAATTACAAAC	300
Qy	301	AAGAAAACTGAAACGCACTTTTAATTGAAACAAACATGCTTTTCTAAATCTTGAAGAAGT	360
Db	301	AAGAAAACTGAAACGCACTTTTAATTGAAACAAACATGCTTTTCTAAATCTTGAAGAAGT	360
Qy	361	TTGGAAGAAATGGATGATGAAGAAGTGAAGAAATGTGCGATGAATATCCATAAAAAAC	420
Db	361	TTGGAAGAAATGGATGATGAAGAAGTGAAGAAATGTGCGATGAATATCCATAAAAAAC	420
Qy	421	ACGGACAATTTAAACAGACAAGCCTTTAAGCGGGCGACTTTAAACGCTGTTTAAATGGG	480
Db	421	ACGGACAATTTAAACAGACAAGCCTTTAAGCGGGCGACTTTAAACGCTGTTTAAATGGG	480
Qy	481	GGTTTTAAATCTTATCAATTAAGTGTCAATGCTGCGGAATCGGTCGCAAAAACCATTTCTA	540
Db	481	GGTTTTAAATCTTATCAATTAAGTGTCAATGCTGCGGAATCGGTCGCAAAAACCATTTCTA	540
Qy	541	GGGCGTGGTTTTATCGCTTGGCGGCAATCAGGTGCTTACAAAGAACTCTCAGAGCTTTTAAACA	600
Db	541	GGGCGTGGTTTTATCGCTTGGCGGCAATCAGGTGCTTACAAAGAACTCTCAGAGCTTTTAAACA	600
Qy	601	GGTCCTGTTGGCTGGATCAATACAGGCGTATGACAGCGATTGATATTTGCAGGGCCGGCT	660
Db	601	GGTCCTGTTGGCTGGATCAATACAGGCGTATGACAGCGATTGATATTTGCAGGGCCGGCT	660
Qy	661	TATAGGTAACCAACCGGCAATGATGTTGGTTCACATTTACGCTTAAACAAACACAGCAA	720
Db	661	TATAGGTAACCAACCGGCAATGATGTTGGTTCACATTTACGCTTAAACAAACACAGCAA	720
Qy	721	GCCAAATGAGATAGAGTGGTTGCAATAGAAATCCATT	759
Db	721	GCCAAATGAGATAGAGTGGTTGCAATAGAAATCCATT	759

RESULT 4
 PCT-US01-48392-43
 ; Sequence 43, Application PC/TUS0148392
 ; GENERAL INFORMATION:
 ; APPLICANT: AntexBiologicals, Inc.
 ; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
 ; TITLE OF INVENTION: thereof
 ; FILE REFERENCE: 7969-091-228
 ; CURRENT APPLICATION NUMBER: PCT/US01/48392
 ; CURRENT FILING DATE: 2001-12-07
 ; PRIOR APPLICATION NUMBER: 09/732,091
 ; PRIOR FILING DATE: 2000-12-07
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 43
 ; LENGTH: 795
 ; TYPE: DNA
 ; ORGANISM: Helicobacter sp.
 PCT-US01-48392-43

```

RESULT 4
PCT-US01-48392-43
; Sequence 43, Application PC/TUS0148392
; GENERAL INFORMATION:
; APPLICANT: AntexBiologicals, Inc.
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-091-228
; CURRENT APPLICATION NUMBER: PCT/US01/48392
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Helicobacter sp.
PCT-US01-48392-43

```

	Query Match	100.0%	Score 759;	DB 1;	Length 795;
	Best Local Similarity	100.0%;	Pred. NO. 1.4e-221;		
	Matches 759;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	ATGGCATACAAATATGATAGAGACTTCGGAATTTTAAAGCAATCTGGAAATCTAGTGATTTA	60		
Db	37	ATGGCATACAAATATGATAGAGACTTCGGAATTTTAAAGCAATCTGGAAATCTAGTGATTTA	96		
Qy	61	TTTGATTTGTTTGAGGTGTCCTGTTTTGGTAAAGACGGCGGAAAAAAGACACAATGAAAAA	120		


```
Query Match      100.0%; Score 759; DB 50; Length 795;
Best Local Similarity 100.0%; Pred. No. 1.4e-221;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTATTTA 60
DB 37 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTATTTA 96
QY 61 TTGGATTTGTTGAGTCTCTGTTTGTGTAAGACGGGAAAGACACAAATGAAGAA 120
DB 97 TTGGATTTGTTGAGTCTCTGTTTGTGTAAGACGGGAAAGACACAAATGAAGAA 156
QY 121 CTGACAGCTCCATAGAAATCAAAAAGGATGCGGATGATTACGCTAAATACGAGAA 180
DB 157 CTGACAGCTCCATAGAAATCAAAAAGGATGCGGATGATTACGCTAAATACGAGAA 216
QY 181 ATCGCTGAAGAGTTGCAATATATATGAGGCAATAGTTTTCGAGTTTCATTAAGGCGAA 240
DB 217 ATCGCTGAAGAGTTGCAATATATATGAGGCAATAGTTTTCGAGTTTCATTAAGGCGAA 276
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTTGCGGATTAATTAAGGTCAATTACAC 300
DB 277 GGAGTCTTATACAAAGAGATTTTATGCGATGTTGCGGATTAATTAAGGTCAATTACAC 336
QY 301 AAGAAACTGAAACGACTTTTAAATGCAACAAACATGCTTTCTAAATCTTGAAGAGAGT 360
DB 337 AAGAAACTGAAACGACTTTTAAATGCAACAAACATGCTTTCTAAATCTTGAAGAGAGT 396
QY 361 TTGGAAGAAATGGATGATGAAGAGTGAAGAAATGTGCGATGAATTTCCATTAAGAAAC 420
DB 397 TTGGAAGAAATGGATGATGAAGAGTGAAGAAATGTGCGATGAATTTCCATTAAGAAAC 456
QY 421 ACAGCAATTTAAACAGACAGCCTTAAGCGCGGCGACTTTAAACGCTGTTTAAATGGGG 480
DB 457 ACAGCAATTTAAACAGACAGCCTTAAGCGCGGCGACTTTAAACGCTGTTTAAATGGGG 516
QY 481 GGTCTTAAATCTTATCAATAGTCTGTCATTTGCGAATGCGGTGCGCAAAACCACTTCTA 540
DB 517 GGTCTTAAATCTTATCAATAGTCTGTCATTTGCGAATGCGGTGCGCAAAACCACTTCTA 576
QY 541 GGGCGTGTGTTTATCGCTGCGGCAATCAGGTGCTTTCAAGAACTCTGAGCTTTTAAACA 600
DB 577 GGGCGTGTGTTTATCGCTGCGGCAATCAGGTGCTTTCAAGAACTCTGAGCTTTTAAACA 636
QY 601 GGTCTGTGTGCTGATCATTTACAGGCGTATGGAACAGCGATTTGATATTGAGGGCGGCT 660
DB 637 GGTCTGTGTGCTGATCATTTACAGGCGTATGGAACAGCGATTTGATATTGAGGGCGGCT 696
QY 661 TATAGGGTAAACCATACCGGCATGCTTGTGTTGCGCACTTTACGCTTAAACACAGCA 720
DB 697 TATAGGGTAAACCATACCGGCATGCTTGTGTTGCGCACTTTACGCTTAAACACAGCA 756
QY 721 GCCAATGAGATAGAAAGTCTGTTGCAATAGAAATCCATT 759
DB 757 GCCAATGAGATAGAAAGTCTGTTGCAATAGAAATCCATT 795
```

RESULT 7

```
PCT-US96-18542-91
; Sequence 91, Application PC/TUS9618542
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 268
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
```

```
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER:
OPERATING SYSTEM:
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/18542
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/561,469
FILING DATE: 17-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-001CPPC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 762 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: 1...762
PCT-US96-18542-91
```

```
Query Match      95.8%; Score 727; DB 3; Length 762;
Best Local Similarity 97.4%; Pred. No. 9.6e-212;
Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTATTTA 60
DB 1 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTATTTA 60
QY 61 TTGGATTTGTTGAGTCTCTGTTTGTGTAAGACGGGAAAGACACAAATGAAGAA 120
DB 61 TTGGATTTGTTGAGTCTCTGTTTGTGTAAGACGGGAAAGACACAAATGAAGAA 120
QY 121 CTGACAGCTCCATAGAAATCAAAAAGGATGCGGATGATTACGCTAAATACGCAAGAA 180
DB 121 CTGACAGCTCCATAGAAATCAAAAAGGATGCGGATGATTACGCTAAATACGCAAGAA 180
QY 181 ATCGCTGAAGAGTTGCAATATCTATGAGGAGCAATAGTTTTCGAGATTTTAAAGGCGAA 240
DB 181 ATCGCTGAAGAGTTGCAATATCTATGAGGAGCAATAGTTTTCGAGATTTTAAAGGCGAA 240
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTTGCGGATTAATTAAGGTCAATTACAC 300
DB 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTTGCGGATTAATTAAGGTCAATTACAC 300
QY 301 AAGAAACTGAAACGACTTTTAAATGCAACAAACATGCTTTCTAAATCTTGAAGAGAGT 360
DB 301 AAGAAACTGAAACGACTTTTAAATGCAACAAACATGCTTTCTAAATCTTGAAGAGAGT 360
QY 361 TTGGAAGAAATGGATGATGAAGAGTGAAGAAATGCGATGAATTTATCCATTAAGAAAC 420
DB 361 CTAGAAGAAATGGATGATGAAGAGTGAAGAAATGCGATGAATTTATCCATTAAGAAAC 420
QY 421 ACAGCAATTTAAACAGACAGCCTTAAAGCGGCGACTTTAAACGCTGTTTAAATGGGG 480
DB 421 ACAGCAATTTAAACAGACAGCCTTAAAGCGGCGACTTTAAACGCTGTTTAAATGGGG 480
QY 481 GGTCTTAAATCTTATCAATAGTCTGTCATTTGCGAATGCGGTGCGCAAAACCACTTCTA 540
```

Db 481 GGCTTTAAATCTTATCAATAGCTGCTCATTTGTCGAATCGGTTGCAAAACCAATCTTA 540
QY 541 GGGCGTGTATTCGCTTGGGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAAACA 600
Db 541 GGGCGTGTATTCGCTTGGGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAAACA 600
QY 601 GGTCTCTTGGCTGGATCATTTACAGGGGTATGGACGAGTTGATATTGCAAGGCGCGCT 660
Db 601 GGCCTCTTGGCTGGATCATTTACAGGGGTATGGACGAGTTGATATTGCAAGGCGCGCT 660
QY 661 TATAGGTAAACCAATACCGGCAATGCTGTGTCGCACTTTACGCTTAAACACAGCAA 720
Db 661 TATAGGTAAACCAATACCGGCAATGCTGTGTCGCACTTTACGCTTAAACACAGCAA 720
QY 721 GCCAATGAAGATAGAGTCTGTCGAATAGAAATCCATT 759
Db 721 GCCAATGAAGATAGAGTCTGTCGAATAGAAATCCGTT 759

RESULT 8

US-08-561-469A-971
; Sequence 971, Application US/08561469A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 994
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/561,469A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 971:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 762 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
US-08-561-469A-971

Query Match 95.8%; Score 727; DB 14; Length 762;
Best Local Similarity 97.4%; Pred. No. 9.6e-212;
Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1 ATGGCATAACAATATGATAGAGACTTGGAAATTTTAAAGCAATCGGAATCTAGTGATTTA 60
Db 1 ATGGCATAACAATATGATAGAGACTTGGAAATTTTAAAGCAATCGGAATCTAGTGATTTA 60

QY 61 TTGATTTGTTGAGGTGCTTTGTTTGGTAAAGACGGCGAAAAAGACACAATGAAAA 120
Db 61 TTGATTTGTTGAGGTGCTTTGTTTGGTAAAGACGGCGAAAAAGACACAATGAAAA 120
QY 121 CTGACAGCTCATAGATAACAAAGGCAATGCGGATGATTACGCTAAATACGACAGAA 180
Db 121 CTCAAGCTCATAGATAACAAAGGCAATGCGGATGATTACGCTAAATACGACAGAA 180
QY 181 ATCGCTGAAGAGTTGCAATACTATGGGAGCAATAGTTTTCGCGAGTTTCAATTAAGGCGAA 240
Db 181 ATCGCTGAAGAGTTGCAATACTATGGGAGCAATAGTTTTCGCGAGTTTCAATTAAGGCGAA 240
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGATAAAATTAAGGTCAATTAACAAC 300
Db 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGATAAAATTAAGGTCAATTAACAAC 300
QY 301 AAGAAACTGAAACGACTTTTAAATGAAACAAACATGCTTTTAAATCTTAAAGAAAGT 360
Db 301 AAGAAACTGAAACGACTTTTAAATGAAACAAACATGCTTTTAAATCTTAAAGAAAGT 360
QY 361 TTGGAAGAAATGCGATGATGAAGAGTGAAGAAATGTGCGATGAATTTCCATATAAAAC 420
Db 361 TTGGAAGAAATGCGATGATGAAGAGTGAAGAAATGTGCGATGAATTTCCATATAAAAC 420
QY 421 ACGGCAATTTTAAACAGACAAAGCTTAAAGCGCGGACCTTAAACGCTGTTTAAATGGGG 480
Db 421 ACGGCAATTTTAAACAGACAAAGCTTAAAGCGCGGACCTTAAACGCTGTTTAAATGGGG 480
QY 481 GGTTTTAAATCTTATCAATAGCTGCTCATTTGTCGAATGCGGTCCGCAAAACCAATCTTA 540
Db 481 GGTTTTAAATCTTATCAATAGCTGCTCATTTGTCGAATGCGGTCCGCAAAACCAATCTTA 540
QY 541 GGGCGTGTGTTATCGCTTGGCGGCAATCAGGTGCTTACAAGAACTCTGAGCTTTTAAACA 600
Db 541 GGGCGTGTGTTATCGCTTGGCGGCAATCAGGTGCTTACAAGAACTCTGAGCTTTTAAACA 600
QY 601 GGTCTGTGTCGTGATCATTTACAGCGTATGACAGCGGATGATATTCAGGCGCGGCT 660
Db 601 GGTCTGTGTCGTGATCATTTACAGCGTATGACAGCGGATGATATTCAGGCGCGGCT 660
QY 661 TATAGGTAACCATACCGGCAATGCTTGGTTCGCAATAGAAATCCATT 759
Db 661 TATAGGTAACCATACCGGCAATGCTTGGTTCGCAATAGAAATCCGTT 759

RESULT 9

US-08-761-184-48
; Sequence 48, Application US/08761184
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH ET AL
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; NUMBER OF SEQUENCES: 1810
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,184


```
/ TELEPHONE: (617)227-7400
/ TELEFAX: (617)227-5941
/ INFORMATION FOR SEQ ID NO: 649:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 762 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Helicobacter pylori
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 1...762
/ US-08-761-184-649

Query Match          95.8%; Score 727; DB 16; Length 762;
Best Local Similarity 97.4%; Pred. No. 9.6e-212;
Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTATT 60
DB 1 ATGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTATT 60
QY 61 TTGGATTGTTTGGAGTCTGCTTTTGGTAAAGCGCGGAAAGACACAAATGAAAA 120
DB 61 TTGGATTGTTTGGAGTCTGCTTTTGGTAAAGCGCGGAAAGACACAAATGAAAA 120
QY 121 CTGACCAAGCTCCATAGAAATACAAAGGATCGCGATGATTACGCTAAATAGCGAAAGA 180
DB 121 CTCAAGCTCCATAGAAATACAAAGGATCGCGATGATTACGCTAAATAGCGAAAGA 180
QY 181 ATCGCTGAAGGTTGCAATCTACTATGGGAGCAATAGTTTTCGAGTTTCATTAAGGCGAA 240
DB 181 ATCGCTGAAGGTTGCAATCTACTATGGGAGCAATAGTTTTCGAGTTTCATTAAGGCGAA 240
QY 241 GGAGTCTTATCAAAAGAGATTTTATGCGATGTGCGATTAATTAAGGTCATTAACAC 300
DB 241 GGAGTCTTATCAAAAGAGATTTTATGCGATGTGCGATTAATTAAGGTCATTAACAC 300
QY 301 AAGAAACTGAAAGCACTTTAATTTGAACAAACATGCTTTCTAAATCTTGAAGAAGT 360
DB 301 AAGAAACTGAAAGCACTTTAATTTGAACAAACATGCTTTCTAAATCTTGAAGAAGT 360
QY 361 TTGAAGAAATGGATGATGAAGAAGTGAAGAAGTGTGCGATGAATTAATCCATAAAAAAC 420
DB 361 CTAGAAGAAATGGATGATGAAGAAGTGAAGAAGTGTGCGATGAATTAATCCATAAAAAAC 420
QY 421 ACGGCAATTTAAACAGCAAGCTTTAAGCGCGGCACTTTAAACGCTGTTTAAATGGGG 480
DB 421 ACGGCAATTTAAACAGCAAGCTTTAAGCGCGGCACTTTAAACGCTGTTTAAATGGGG 480
QY 481 GGTTTTAATCTTATCAATAGCTGCTCAATGTCGAATGCGGTGCGTAAACACCACTTCTA 540
DB 481 GGTTTTAATCTTATCAATAGCTGCTCAATGTCGAATGCGGTGCGTAAACACCACTTCTA 540
QY 541 GGGGCTGTTTATCGCTTGGGGCAATCAGGTGCTTTACAAGAACTCTGAGCTTTTAAACA 600
DB 541 GGGGCTGTTTATCGCTTGGGGCAATCAGGTGCTTTACAAGAACTCTGAGCTTTTAAACA 600
QY 601 GGTCTGTTGGTGGATCAATTAAGGCGTATGGACAGCAATGATATTGACGGGCGGCT 660
DB 601 GGTCTGTTGGTGGATCAATTAAGGCGTATGGACAGCAATGATATTGACGGGCGGCT 660
QY 661 TATAGGGTAACATACCGGCATGCAATGTTGGTGGCACTTTACGCTTAAACACAGCAA 720
DB 661 TATAGGGTAACATACCGGCATGCAATGTTGGTGGCACTTTACGCTTAAACACAGCAA 720
QY 721 GCCAATGAGATAGAGAGTCTGTCAAATAGAAATCCATT 759
DB 721 GCCAATGAGATAGAGAGTCTGTCAAATAGAAATCCATT 759
```

RESULT 11

```
US-08-821-931-48
; Sequence 48, Application US/08821931
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH ET AL
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; NUMBER OF SEQUENCES: 1810
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,931
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/561,469
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,405
; FILING DATE: 01-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/660,742
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/736,791
; FILING DATE: 25-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/739,150
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/761,184
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy B.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001CP9CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 762 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...762
; US-08-821-931-48
```

```
Query Match          95.8%; Score 727; DB 17; Length 762;
Best Local Similarity 97.4%; Pred. No. 9.6e-212;
Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```
QY 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
Db |||
QY 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
Db |||
QY 61 TTGGATTTGTTGAGTGCTGTTGTTTGGTAAAGCGGCGGAAAAAGACACAAATGAAAAA 120
Db |||
QY 61 TTGGATTTGTTGAGTGCTGTTGTTTGGTAAAGCGGCGGAAAAAGACACAAATGAAAAA 120
Db |||
QY 121 CTGACAGCTCCATAGATACAAAAGGCGATGCGGATGATTACGTAATACGCGAGAAAG 180
Db |||
QY 121 CTGACAGCTCCATAGATACAAAAGGCGATGCGGATGATTACGTAATACGCGAGAAAG 180
Db |||
QY 181 ATCGCTGAAGAGTTGCAATACATATGCGGAGCAATAGTTTTCGAGTTTCAATTAAGCGGAA 240
Db |||
QY 181 ATCGCTGAAGAGTTGCAATACATATGCGGAGCAATAGTTTTCGAGTTTCAATTAAGCGGAA 240
Db |||
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 300
Db |||
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 300
Db |||
QY 301 AAGAAACCTGAACAGCTTTAATGGAACAAACATGCTTTCTAAATCTTAGAAGAGCT 360
Db |||
QY 301 AAGAAACCTGAACAGCTTTAATGGAACAAACATGCTTTCTAAATCTTAGAAGAGCT 360
Db |||
QY 361 TTGGAAGAAATGATGATGAAGAGTGAAGAAATGTCGATGAATATCCATATAAAAC 420
Db |||
QY 361 CTAGAAAGAAATGATGATGAAGAGTGAAGAAATGTCGATGAATATCCATATAAAAC 420
Db |||
QY 421 ACGGCAATTTAAACAGACAGCTTTAAGCGCGGAGCTTTAAGCGCTGTTTAAATGGGG 480
Db |||
QY 421 ACGGCAATTTAAACAGACAGCTTTAAGCGCGGAGCTTTAAGCGCTGTTTAAATGGGG 480
Db |||
QY 481 GGTTTTAAATCTTATCAATAGCTGCTGATGTTGCGAATGCGGTCGCAAAACCATTTCTA 540
Db |||
QY 481 GGTTTTAAATCTTATCAATAGCTGCTGATGTTGCGAATGCGGTCGCAAAACCATTTCTA 540
Db |||
QY 541 GGGCGTGTTTATCGCTGCGGCAATCAGGTGCTTCAAGAACTCTGAGCTTTTAAAC 600
Db |||
QY 541 GGGCGTGTTTATCGCTGCGGCAATCAGGTGCTTCAAGAACTCTGAGCTTTTAAAC 600
Db |||
QY 601 GGTCTGTTGGCTGATCATACAGCGCTATGACAGCGATTTGATATTCAGCGCGCGCT 660
Db |||
QY 601 GGTCTGTTGGCTGATCATACAGCGCTATGACAGCGATTTGATATTCAGCGCGCGCT 660
Db |||
QY 661 TATAGGTAAACATACCGGATGCAATGTTGTTGCGCACTTTACGCTAAAAACACAGCAA 720
Db |||
QY 661 TATAGGTAAACATACCGGATGCAATGTTGTTGCGCACTTTACGCTAAAAACACAGCAA 720
Db |||
QY 721 GCCAATGAGATAGAAAGTCTGTTGCAATAGATCCATT 759
Db |||
QY 721 GCCAATGAGATAGAAAGTCTGTTGCAATAGATCCATT 759
Db |||
```

RESULT 12

```
US-08-821-931-649
; Sequence 649, Application US/08821931
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH ET AL
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; NUMBER OF SEQUENCES: 1810
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: ,
```

```
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,931
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/561,469
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,405
; FILING DATE: 01-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/660,742
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/736,791
; FILING DATE: 25-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/739,150
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/761,184
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001CP9CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 649:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 762 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...762
; US-08-821-931-649
```

```
Query Match 95.8%; Score 727; DB 17; Length 762;
Best Local Similarity 97.4%; Pred. No. 9.6e-212;
Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
Db |||
QY 1 ATGGCATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
Db |||
QY 61 TTGGATTTGTTGAGTGCTGTTGTTTGGTAAAGCGGCGGAAAAAGACACAAATGAAAAA 120
Db |||
QY 61 TTGGATTTGTTGAGTGCTGTTGTTTGGTAAAGCGGCGGAAAAAGACACAAATGAAAAA 120
Db |||
QY 121 CTGACAGCTCCATAGATACAAAAGGCGATGCGGATGATTACGTAATACGCGAGAAAG 180
Db |||
QY 121 CTGACAGCTCCATAGATACAAAAGGCGATGCGGATGATTACGTAATACGCGAGAAAG 180
Db |||
QY 181 ATCGCTGAAGAGTTGCAATACATATGCGGAGCAATAGTTTTCGAGTTTCAATTAAGCGGAA 240
Db |||
QY 181 ATCGCTGAAGAGTTGCAATACATATGCGGAGCAATAGTTTTCGAGTTTCAATTAAGCGGAA 240
Db |||
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 300
Db |||
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 300
Db |||
```

```
QY 301 AAGAAACTGAACGACTTTAAATTGAACAAAAACATGCTTTCTAAATCTTAGAAGAAGT 360
Db |||||
QY 301 AAGAAACTGAACGACTTTAAATTGAACAAAAACATGCTTTCTAAATCTTAGAAGAAGT 360
Db |||||
QY 361 TTGGAAGAAATGGATGATGAAGAAGTGAAGAAATGTGCGATGAATATCCATAAAAAAC 420
Db |||||
QY 361 CTAGAAGAAATGGATGATGAAGAAGTGAAGAAATGTGCGATGAATATCCATAAAAAAC 420
Db |||||
QY 421 ACGGACAAATTAACACAGCAAGCTTTAAGCGCGGCACTTTAAGCGCTTTAAATGGG 480
Db |||||
QY 421 ACGGACAAATTTGAACAGCAAGCTTTAAGCGCGGCACTTTAAGCGCTTTAAATGGG 480
Db |||||
QY 481 GGTTTAAATCTTAATCAATAGCTGCTCATTTGCGAATGCGTGCAGAAAAACCAATCTTA 540
Db |||||
QY 481 GGTTTAAATCTTAATCAATAGCTGCTCATTTGCGAATGCGTGCAGAAAAACCAATCTTA 540
Db |||||
QY 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAA 600
Db |||||
QY 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGTGTCTTACAAGAACTCTGAGCTTTTAA 600
Db |||||
QY 601 GGTCTGTGTGCTGGATCATTAAGCGGTATGAGCAGCGATGATATGCGAGGCGCGCT 660
Db |||||
QY 601 GGGCGTGTGTGCTGGATCATTAAGCGGTATGAGCAGCGATGATATGCGAGGCGCGCT 660
Db |||||
QY 661 TATAGGGAATACCATACCGGATGCTTGTGCGCAATGAGTGTGCGCAATGAGTGTGCG 720
Db |||||
QY 661 TATAGGGAATACCATACCGGATGCTTGTGCGCAATGAGTGTGCGCAATGAGTGTGCG 720
Db |||||
QY 721 GCCAATGGAGATAAGAAGTGTGCGCAATGAGTGTGCGCAATGAGTGTGCGCAATGAGT 759
Db |||||
QY 721 GCCAATGGAGATAAGAAGTGTGCGCAATGAGTGTGCGCAATGAGTGTGCGCAATGAGT 759
Db |||||
```

RESULT 13

US-08-993-002A-4400
; Sequence 4400, Application US/08993002A

GENERAL INFORMATION:

; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031

CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/993.002A
; FILING DATE:

PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 4400:

SEQUENCE CHARACTERISTICS:

; LENGTH: 762 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Helicobacter pylori

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1...762

; US-08-993-002A-4400

Query Match

Best Local Similarity 95.8%; Score 727; DB 18; Length 762;

Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
QY 1 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGAAATCTAGTGAATTTA 60
Db |||||
QY 1 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGAAATCTAGTGAATTTA 60
Db |||||
QY 61 TTGGATTTTGTGAGGTGCTTGTGTTTGGTAAAGACGGCAAAAAAGACACAAATGAAAAA 120
Db |||||
QY 61 TTGGATTTTGTGAGGTGCTTGTGTTTGGTAAAGACGGCAAAAAAGACACAAATGAAAAA 120
Db |||||
QY 121 CTGACCACTCCATAGAATACAAAAGGCATGCGATGATTACGCTAAATACGACAGAAAGA 180
Db |||||
QY 121 CTCACAAGCTCCATAGAATACAAAAGGCATGCGATGATTACGCTAAATACGACAGAAAGA 180
Db |||||
QY 181 ATCGCTGAAGAGTTGCCAATACCTATGGAGCAATAGTTTTCGCGAGTTTCATTAAAGGCG 240
Db |||||
QY 181 ATCGCTGAAGAGTTGCCAATACCTATGGAGCAATAGTTTTCGCGAGTTTCATTAAAGGCG 240
Db |||||
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGATAAATAAAGGTCAATTTACAAC 300
Db |||||
QY 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGATAAATAAAGGTCAATTTACAAC 300
Db |||||
QY 301 AAGAAAACGTGAACGACTTTAAATGAACAAAAACATGCTTTCTAAAAATCTTAGAAAGAGT 360
Db |||||
QY 301 AAGAAAACGTGAACGACTTTAAATGAACAAAAACATGCTTTCTAAAAATCTTAGAAAGAGT 360
Db |||||
QY 361 TTGGAAGAAATGGATGATGAAGAAGTGAAGAAATGTGCGATCAATATCCATAAAAAAC 420
Db |||||
QY 361 CTAGAAGAAATGGATGATGAAGAAGTGAAGAAATGTGCGATCAATATCCATAAAAAAC 420
Db |||||
QY 421 ACGGACAAATTTAAACACAGCAAGCTTTAAGCGCGGCACTTTAAGCGCTTTTAAAAATGGGG 480
Db |||||
QY 421 ACGGACAAATTTAAACACAGCAAGCTTTAAGCGCGGCACTTTAAGCGCTTTTAAAAATGGGG 480
Db |||||
QY 481 GGTGTTAAATCTTATCAATTAGCTGTCAATTTGTCGAATGCGGTGCGCAAAACCAATTTCTA 540
Db |||||
QY 481 GGTGTTAAATCTTATCAATTAGCTGTCAATTTGTCGAATGCGGTGCGCAAAACCAATTTCTA 540
Db |||||
QY 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTTACAAGAACTCTGAGCTTTTAAACA 600
Db |||||
QY 541 GGGCGTGGTTTATCGCTTGGCGGCAATCAGGTGCTTTACAAGAACTCTGAGCTTTTAAACA 600
Db |||||
QY 601 GGTCTGTGTGCTGGATCATTAACAGCGGTATGAGCAGCGATGATATGCGAGGCGCGCT 660
Db |||||
QY 601 GGTCTGTGTGCTGGATCATTAACAGCGGTATGAGCAGCGATGATATGCGAGGCGCGCT 660
Db |||||
QY 661 TATAGGTAACCATACCGGCACTTTGTTGCGCACTTTACGCTTAAACACACAGCAA 720
Db |||||
QY 661 TATAGGTAACCATACCGGCACTTTGTTGCGCACTTTACGCTTAAACACACAGCAA 720
Db |||||
QY 721 GCCAATGGAGATAAGAAGTGTGTTGCGCAATGAGTGTGCGCAATGAGTGTGCGCAATGAGT 759
Db |||||
QY 721 GCCAATGGAGATAAGAAGTGTGTTGCGCAATGAGTGTGCGCAATGAGTGTGCGCAATGAGT 759
Db |||||
```

RESULT 14

US-08-993-002A-4401

; Sequence 4401, Application US/08993002A

; GENERAL INFORMATION:

; APPLICANT: DOUGLAS SMITH et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: Windows NT 4.0

SOFTWARE: UNIX

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,002A

FILING DATE:

PRIOR APPLICATION NUMBER:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: GTN-018

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 4401:

SEQUENCE CHARACTERISTICS:

LENGTH: 762 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Helicobacter pylori

FEATURE:

NAME/KEY: misc_feature

LOCATION: 1...762

US-08-993-002A-4401

Query Match 95.8%; Score 727; DB 18; Length 762;

Best Local Similarity 97.4%; Pred. No. 9.6e-212;

Matches 739; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTCGAATCTAGTGATTTA 60

DB 1 ATGGCATCAATATGATAGAGACTTGGAAATTTTAAAGCAATTCGAATCTAGTGATTTA 60

QY 61 TTGGATTGTTTGAAGTCTGTTTGGTAAAGACGCGGAAAAAGACACAATGAAAAA 120

DB 61 TTGGATTGTTTGAAGTCTGTTTGGTAAAGACGCGGAAAAAGACACAATGAAAAA 120

QY 121 CTGACAGCTCCATAGATAAACAAGGATCGCGATGATTAACGTAATATAGCAGAAAGA 180

DB 121 CTGACAGCTCCATAGATAAACAAGGATCGCGATGATTAACGTAATATAGCAGAAAGA 180

QY 181 ATCGCTGAAGTTGCAATATGATGGAGCAATAGTTTGGAGTTTCATTAAAGCGCAA 240

DB 181 ATCGCTGAAGTTGCAATATGATGGAGCAATAGTTTGGAGTTTCATTAAAGCGCAA 240

QY 241 GGAGTCTTATCAAGAGATTTTATGCGATGTTGCGGATAAATTAAGGTCAATTAACAAC 300

DB 241 GGAGTCTTATCAAGAGATTTTATGCGATGTTGCGGATAAATTAAGGTCAATTAACAAC 300

QY 301 AAGAAAACTGAACGACTTTTAATTGAACAAACATGCTTTCTAAATCTTGAAGAAGAGT 360

DB 301 AAGAAAACTGAACGACTTTTAATTGAACAAACATGCTTTCTAAATCTTGAAGAAGAGT 360

QY 361 TTGGAGAAATGGATGATGAAGAGTGAAGAAATGTCGGATGAATATTCATTAATAAAC 420

DB 361 CTAGAGAAATGGATGATGAAGAGTGAAGAAATGTCGGATGAATATTCATTAATAAAC 420

QY 421 ACGGCAATTTAAACAGACAGAGCTTTAAGCGCGGACCTTTAAGCGCTGTTTAAATGGGG 480

DB 421 ACGGCAATTTAAGACAGACAGAGCTTTAAGCGCGGACCTTTAAGCGCTGTTTAAATGGGA 480

QY 481 GGTCTTAAATCTTATCAATTTAGCTGTCATTGTTGCGAATCGGTCGCAAAACCATTTCTA 540

DB 481 GGTCTTAAATCTTATCAATTTAGCTGTCATTGTTGCGAATCGGTCGCAAAACCATTTCTA 540

QY 541 GGGCGTGGTTTATCGCTTGGGGCAATCAGGTCTTACAGAACTCTGAGCTTTTAAACA 600

DB 541 GGGCGTGGTTTATCGCTTGGGGCAATCAGGTCTTACAGAACTCTGAGCTTTTAAACA 600

QY 601 GGTCTTGGTGGTGGATCAATACAGCGGTATGACAGCGATTTGATATTGACGGCCGGCT 660

DB 601 GGTCTTGGTGGTGGATCAATACAGCGGTATGACAGCGATTTGATATTGACGGCCGGCT 660

QY 661 TATAGGTTAACCATACCGGCATGTCATTGTTGCGCACTTTACGCTTAAACACAGCAA 720

DB 661 TATAGGTTAACCATACCGGCATGTCATTGTTGCGCACTTTACGCTTAAACACAGCAA 720

QY 721 GCCAATGAGATAAGAGTGGTTGCGAAATGAATCCATT 759

DB 721 GCCAATGAGATAAGAGTGGTTGCGAAATGAATCCATT 759

RESULT 15

US-10-335-977-4400

Sequence 4400, Application US/10335977

GENERAL INFORMATION:

APPLICANT: DOUGLAS SMITH et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

RELATING TO HELICOBACTER PYLORI FOR

DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: Windows NT 4.0

SOFTWARE: UNIX

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/335,977

FILING DATE: 30-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/993,002

FILING DATE: 17-DEC-1997

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: GTN-018

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 4400:

SEQUENCE CHARACTERISTICS:

LENGTH: 762 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Helicobacter pylori

```
;
;
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...762
; SEQUENCE DESCRIPTION: SEQ ID NO: 4400:
US-10-335-977-4400

Query Match          95.8%; Score 727; DB 49; Length 762;
Best Local Similarity 97.4%; Pred. No. 9.6e-212; Indels 0; Gaps 0;
Matches 739; Conservative 0; Mismatches 20;

Qy 1 ATGCGATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60
Db 1 ATGCGATACAAATATGATAGAGACTTGGAAATTTTAAAGCAATTTGGAATCTAGTGATTTA 60

Qy 61 TTGGATTGTTTGGAGTGCTTGTGTTTGGTAAAGACGCGGAAAGACACAAATGCAAAA 120
Db 61 TTGGATTGTTTGGAGTGCTTGTGTTTGGTAAAGACGCGGAAAGACACAAATGCAAAA 120

Qy 121 CTGACCAAGCTCCATAGAAATCAAAAGGATGGCGATGATTACGCTAAATACGCAGAAAGA 180
Db 121 CTGACCAAGCTCCATAGAAATCAAAAGGATGGCGATGATTACGCTAAATACGCAGAAAGA 180

Qy 181 ATCGCTCAAGAGTTGCAATACTATGGAGCAATAGTTTTCGAGTTTCATTAAAGGCGAA 240
Db 181 ATCGCTCAAGAGTTGCAATACTATGGAGCAATAGTTTTCGAGTTTCATTAAAGGCGAA 240

Qy 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGGATATTAATTAAGTCAATTTACAA 300
Db 241 GGAGTCTTATACAAAGAGATTTTATGCGATGTGCGGATATTAATTAAGTCAATTTACAA 300

Qy 301 AAGAAACTGAAAGCACTTTAAATTAAGCAAAACATGCTTCTAAATCTTAGAAGAGAGT 360
Db 301 AAGAAACTGAAAGCACTTTAAATTAAGCAAAACATGCTTCTAAATCTTAGAAGAGAGC 360

Qy 361 TTGAAGAAATGGATGATGAAGAGTGAAGAAATGTCGATGAATTTATCCATAAAAC 420
Db 361 CTAGAAGAAATGGATGATGAAGAGTGAAGAAATGTCGATGAATTTATCCATAAAAC 420

Qy 421 ACGGACAAATTTAAACAGACAGCCCTTAAGCGCGGCACTTTAAACGCTTTTAAATGGG 480
Db 421 ACGGACAAATTTAAACAGACAGCCCTTAAGCGCGGCACTTTAAACGCTTTTAAATGGG 480

Qy 481 GGTTTTAAATCTTATCAATTAAGTCTGTCATTTGCGAATGCGTCCGAAACCATTTCTA 540
Db 481 GGCTTTAAATCTTATCAATTAAGTCTGTCATTTGCGAATGCGTTCGAAACCATTTCTA 540

Qy 541 GGGCGTGTTTTATCGCTTGGCGCAATCAGGTGCTTCAAGAACTCTGAGCTTTTAAACA 600
Db 541 GGGCGTGTTTTATCGCTTGGCGCAATCAGGTGCTTCAAGAACTCTGAGCTTTTAAACA 600

Qy 601 GGTCTGTGTGGCTGGATCAATTAAGCGGCTATGGACAGCGATTTGATTAATGCGGGCGGCT 660
Db 601 GGCCTGTGTGGCTGGATCAATTAAGCGGCTATGGACAGCGATTTGATTAATGCGGGCGGCT 660

Qy 661 TATAGGGTAACCATACCGGCAATGATTTGGTTGCACTTTACGCTTAAACACAGCAA 720
Db 661 TATAGGGTAACCATACCGGCAATGATTTGGTTGCACTTTACGCTTAAACACAGCAA 720

Qy 721 GCCAATGAGATAGAAGTGGTTCGAATAGATCCATT 759
Db 721 GCCAATGAGATAGAAGTGGTTCGAATAGATCCGTT 759
```

Search completed: August 3, 2006, 10:55:37
Job time : 5923 secs

Result No.	Score	Query			DB	ID	Description
		Match	Length				
C 1	43.4	5.7	556	6	US-10-363-345B-23587	Sequence 23587, A	
C 2	43.4	5.7	556	6	US-10-363-345B-23588	Sequence 23588, A	
C 3	42.8	5.6	1564	6	US-10-363-345B-8863	Sequence 8863, Ap	
C 4	42.8	5.6	1564	6	US-10-363-345B-8864	Sequence 8864, Ap	
C 5	42.4	5.6	639	7	US-11-486-448-1386	Sequence 1386, Ap	
C 6	42	5.5	587	6	US-10-363-345B-12619	Sequence 12619, A	
C 7	42	5.5	587	6	US-10-363-345B-12620	Sequence 12620, A	
C 8	41.6	5.5	645	7	US-11-486-448-7023	Sequence 7023, Ap	
C 9	40.8	5.4	576	6	US-10-363-345B-40153	Sequence 40153, A	
C 10	40.8	5.4	576	6	US-10-363-345B-40154	Sequence 40154, A	
C 11	40.8	5.4	713	6	US-10-363-345B-6193	Sequence 6193, Ap	
C 12	40.8	5.4	713	6	US-10-363-345B-6194	Sequence 6194, Ap	
C 13	40.6	5.3	786	6	US-10-363-345B-20053	Sequence 20053, A	
C 14	40.6	5.3	786	6	US-10-363-345B-20054	Sequence 20054, A	
C 15	40.6	5.3	822	6	US-10-363-345B-31443	Sequence 31443, A	
C 16	40.6	5.3	822	6	US-10-363-345B-31444	Sequence 31444, A	
C 17	40.6	5.3	822	6	US-10-363-345B-38409	Sequence 38409, A	
C 18	40.6	5.3	822	6	US-10-363-345B-38410	Sequence 38410, A	
C 19	40.2	5.3	580	6	US-10-363-345B-21455	Sequence 21455, A	
C 20	40.2	5.3	580	6	US-10-363-345B-21456	Sequence 21456, A	
C 21	40	5.2	780	7	US-11-416-490-20	Sequence 20, Appl	
C 22	39.6	5.2	1762	6	US-10-363-345B-7529	Sequence 7529, Ap	
C 23	39.6	5.2	1762	6	US-10-363-345B-7530	Sequence 7530, Ap	
C 24	39.4	5.2	523643	1	PCT-US03-41389-308	Sequence 308, App	
C 25	39.2	5.2	778	6	US-10-363-345B-2179	Sequence 2179, Ap	

```
QY 401 ATGAATTATCCATAAAAAACGCGACAATTTAAACAGACAA 441
Db 193 AAAAAACGTAAAAAATAAAAAAATAAAAAAATAAAAAA 153

RESULT 2
US-10-363-345B-23588
; Sequence 23588, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 23588
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 23588
; NAME/KEY: unsure
; LOCATION: (503, 506, 513..514, 517, 520..521, 523..524, 552)
US-10-363-345B-23588

Query Match 5.7%; Score 43.4; DB 6; Length 556;
Best Local Similarity 45.5%; Pred. No. 0.035;
Matches 155; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 101 AAAAAACACACAAATGAAACCTGACCGAGTCCATAGATAACAAAGGCGATGATT 160
Db 64 AAAAAAATAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 123

QY 161 ACGCTAATACGCAAGAAAGTCTGCAAGAGTTGCAATCTATGCGAGCAATAGTTTG 220
Db 124 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 183

QY 221 CGAGTTTCATTAAGCGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGCGATA 280
Db 184 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 243

QY 281 AATTAAAGTCAATTACACAGAAACTGAAACGACTTTTAAATTGAACAAACATGCTTT 340
Db 244 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 303

QY 341 CTAAATCTTAGAAGAGTTTGAAGAAATGGATGATGAAGAGTGAAGAAATGTGCG 400
Db 304 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 363

QY 401 ATGAATTATCCATAAAAAACGCGACAATTTAAACAGACAA 441
Db 364 AAAAAACGTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 404

RESULT 3
US-10-363-345B-8863/c
; Sequence 8863, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
```

```
; SEQ ID NO 8863
; LENGTH: 1564
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 8863
; NAME/KEY: unsure
; LOCATION: (9, 21, 31, 38, 54, 161, 265, 288, 328, 342, 390, 1069, 1108)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1199, 1204, 1211..1212, 1222, 1250, 1280..1281, 1348)
US-10-363-345B-8863

Query Match 5.6%; Score 42.8; DB 6; Length 1564;
Best Local Similarity 45.5%; Pred. No. 0.067;
Matches 152; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 108 ACACAATGAAAACTGACCGAGTCCATAGATAACAAAGGCGATGCGATGATTACGCTAA 167
Db 810 AAAAAATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 751

QY 168 ATACGCAAGAAATCGCTGAAGAGTTGCAATCTATGCGAGCAATAGTTTTCGAGTTT 227
Db 750 AACGAAATTCAAAAACGTAAACGACGTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 691

QY 228 CATTAAAGCGAAGAGTCTTATACAAAGAGATTTTATCGGATGTGCGATGATAAATAA 287
Db 690 CAAATAATACAAAAAATACGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 631

QY 288 GGTCAATTACACAGAAAACTGAAACGACTTTTAAATTGAACAAACATGCTTTCTAAAT 347
Db 630 AACCATAAAAAATAACGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 571

QY 348 CTAGAAAAAGATTTGGAAAGAAATCGATCATCAAGAAAGTGAAGAAATGTCGATGAATT 407
Db 570 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 511

QY 408 ATCCATAAAAAACACGAGCAATTTAAACAGACAA 441
Db 510 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 477

RESULT 4
US-10-363-345B-8864
; Sequence 8864, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 8864
; LENGTH: 1564
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 8864
; NAME/KEY: unsure
; LOCATION: (217, 284..285, 315, 343, 353..354, 361, 366, 457, 496, 575)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1223, 1237, 1277, 1300, 1404, 1511, 1527, 1534, 1544, 1556)
US-10-363-345B-8864

Query Match 5.6%; Score 42.8; DB 6; Length 1564;
```

```
Best Local Similarity 45.5%; Pred. No. 0.067;
Matches 152; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

Qy 108 ACACAAATGAAAGTCCAGCTCCATAGATACAAAAGGCGATGATTCGCTAA 167
Db 755 AAAAATTAATATAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 814
Qy 168 ATACGACAGAAAGATCGCTGAAGAGTTGCAATCTATATGGAGCAATAGTTTTCGAGTTT 227
Db 815 AAACGAAATTACAAACGTAACGACGTAAAAAATAAATAAAGATATAAATAACAGAA 874
Qy 228 CATTAAAGCGAAGAGTCTTATACAAAGAGATTTTATCGGATGTGCGGATAAATAA 287
Db 875 CAAATAATACAAAAATACGAAAAAAGAAAAAATATACATAAATAAATATACCAAT 934
Qy 288 GGTCAATTACACAGAAAGCTGAAACGACTTTAATTGAACAAAAACATGCTTTCTAAAT 347
Db 935 AAACGATAAAAAAATAATCGAAAAATTTATAAAAAAAGAAAAAATAAATAAATAA 994
Qy 348 CTTAGAAAGAGTTTGGAAAGAAATGGATGATGAAGAGTGAAGAAATGTCGATCAAT 407
Db 995 AAAATAAAAAAAGCAAAATTAACGAAACGAAAAACCAAAAAAATAAATAAATAA 1054
Qy 408 ATCCATAAAAAACGCGACAATTTTAAACAGACAA 441
Db 1055 AAATAAAAAAACAACGAAATCAAAAAACGA 1088

RESULT 5
US-11-486-448-1386/c
; Sequence 1386, Application US/11486448
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/11/486,448
; CURRENT FILING DATE: 2006-07-14
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1386
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1040_7
US-11-486-448-1386

Query Match 5.6%; Score 42.4; DB 7; Length 639;
Best Local Similarity 53.7%; Pred. No. 0.067;
Matches 88; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 278 ATAAATTAAGGTCAATTACAAAGAAACCTGAACGACTTTAATTGCAACAAAACATGC 337
Db 595 AAAAATAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 536
Qy 338 TTTCTAAATCTTGAAGAAGTTTGGAAAGAAATGGATGATGAAGAGTGAAGAAATGT 397
Db 535 AACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 476
Qy 398 GCGATGAATTATCCATAAATAACGCGACAATTTAAACAGACAA 441
Db 475 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 432

RESULT 6
US-10-363-345B-12619/c
; Sequence 12619, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 12620
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 12620
; NAME/KEY: unsure
; LOCATION: (27..28, 30..)
```

Query Match

5.5%; Score 42; DB 6; Length 587;

```
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 12619
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 12619
; NAME/KEY: unsure
; LOCATION: (556..557, 559..)
US-10-363-345B-12619

Query Match 5.5%; Score 42; DB 6; Length 587;
Best Local Similarity 45.1%; Pred. No. 0.084;
Matches 156; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

Qy 91 AAAGACGGCGAAAAAGACACATGAATAAAGTCCAGCTCCATAGATACAAAAGCAT 150
Db 418 AAAAAAACGAAAAACGAAAAACACGTAAAAACCGAAAAAATAAATAAATATCAT 359
Qy 151 GCGATGATTACGCTAAATACGAGAAAGAAATCGTGAAGAGTTGCAATCTATGCGAGC 210
Db 358 TAGAAAAAAGCATAAAAACGCAAAACAAAAAAGAAAAAAGAAAAAAGAAAAA 299
Qy 211 AATAGTTTTGCGAGTTTTCATTAAGGCGAAGAGTCTTATACAAAGAGATTTTATGCGAT 270
Db 298 AACGACGTAAAAAAGCAAAAAAAGAAAAAATAAATAAATAAATAAATAAATA 239
Qy 271 GTGTCGATAAATAAAGTCAATTACACAGAAAAAAGTGAAGAGTCTTAAATGAACAA 330
Db 238 AAAAAAATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 179
Qy 331 AACATGCTTTCTAAATCTTAGAAGAAGTTTGGAGAAATGATGATGAAGAGTGA 390
Db 178 AAACGCTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 119
Qy 391 GAAATGTCGATGAATTTATCCATAAATAAATAAATAAATAAATAAATAAATA 436
Db 118 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 73

RESULT 7
US-10-363-345B-12620
; Sequence 12620, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 12620
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 12620
; NAME/KEY: unsure
; LOCATION: (27..28, 30..)
```

Query Match

5.5%; Score 42; DB 6; Length 587;

```
Best Local Similarity 45.1%; Pred. No. 0.084;
Matches 156; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 91 AAGACGGCGAAGACCAATGAAAACTGACGCTCCATAGTAATACAAAAGGCAT 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 AAAAAACGAAACGAAACACGTAACCAACGAAAAAATTTAAATATCAT 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 GCGATGATTTACGCTAAATACGCAAGAAATCGCTGAAGTTGCCAATCTATGGGAGC 210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 TAGCAAAAAACGATAAAAAACGCAAAACCAAAAAACGAAAAACGAAAAAATA 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 AATAGTTTTCGCGATTTTCATTAAGCGGAGAGCTTTATACAAAGAGATTTTATGCCAT 270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 AACCAACGTAATAAAAAACGAAAAAATAAAAAAATAAAAAAATAATAATA 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 GTGTGCGATTAATTAAGGTCATTAACAAGAAAACTGAACGACTTTAATTGAACAA 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 AAAAAAATTATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 331 AACATGCTTTCTAAATCTTAGAAGAGTTTGAAGAAATGGATGATGAAGAGTGA 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 AAAACGCTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 391 GAAATGTGCGATGAATTTATCCATAAAAAACACGACCAATTTAAACA 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-11-486-448-7023/c
; Sequence 7023, Application US/11486448
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/11/486,448
; CURRENT FILING DATE: 2006-07-14
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 7023
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: uC-gsromu33B044d09D1
US-11-486-448-7023

Query Match 5.5%; Score 41.6; DB 7; Length 645;
Best Local Similarity 49.5%; Pred. No. 0.11;
Matches 107; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 233 AAGCGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGCGATATAATTTAAAGGTCA 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 638 AAGTGGAGATTAATAAAGAAAGAAAGAAAGAAATTAAGTGAAGAACAAAAAGGAAT 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 ATTACAAACAGAAAACTGAAACGACTTTAATTGAACAAAAATGCTTTCTAAAAATCTTAG 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 578 AAGAGAAAAAGATAAATTAATGAGTTAATAATAATAATAAGGATGAAAAAATAATTA 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 AAGAAAGTTTGAAGAAATGATGATGAAGAGTGAAGAAATGTGCGATGAATTTATCCA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 518 TATATAAAGAGAAAAAATAAAGTGCAAAAAATAAATAAATAAATAAATAAATAAATA 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 TAAAAACACGCAATTTAAACAGACAGCCTTAA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 TAAAAAATTGAAATGAATAACAAAAAGAAATA 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-10-363-345B-40153/c
; Sequence 40153, Application US/10363345B
```

```
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 40153
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 40153
US-10-363-345B-40153

Query Match 5.4%; Score 40.8; DB 6; Length 576;
Best Local Similarity 53.0%; Pred. No. 0.17; 77; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 0;

QY 278 ATAAATTAAAGTCAATTACAAAGAAACTGAAACGACTTTTAATTGAACAAAAATGC 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567 AAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 TTTCTAAATCTTAGAAGAGTTTGGAAAGAAATGATGATGAAGAGTGAAGAAATGT 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 GAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 GCGATGAATTTATCCATAAAAAACACGACAAATTTAAACAGACAA 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 447 ACGATAAATCGACCGTAAAAAATAAATAAATGTAATAAAAAAATA 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-363-345B-40154
; Sequence 40154, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 40154
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 40154
US-10-363-345B-40154

Query Match 5.4%; Score 40.8; DB 6; Length 576;
Best Local Similarity 53.0%; Pred. No. 0.17; 77; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 0;

QY 278 ATAAATTAAAGTCAATTACAAAGAAACTGAAACGACTTTTAATTGAACAAAAATGC 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 AAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 TTTCTAAATCTTAGAAGAGTTTGGAAAGAAATGATGATGAAGAGTGAAGAAATGT 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 GAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 GCGATGAATTTATCCATAAAAAACACGACAAATTTAAACAGACAA 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 ACGATAAATCGACCGTAAAAAATAAATAAATGTAATAAAAAAATA 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
US-10-363-345B-6193/c
; Sequence 6193, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 6193
; LENGTH: 713
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 6193
US-10-363-345B-6193

Query Match          5.4%; Score 40.8; DB 6; Length 713;
Best Local Similarity 49.1%; Pred. No. 0.18;
Matches 108; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 229 ATTAAGCGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 288
DB 413 AATAAAGGAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 354
QY 289 GTCATTTACAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 348
DB 353 ATAAAGGAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 294
QY 349 TTAGAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 408
DB 293 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 234
QY 409 TCCATTAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 448
DB 233 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 194

RESULT 12
US-10-363-345B-6194
; Sequence 6194, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 6194
; LENGTH: 713
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 6194
US-10-363-345B-6194

Query Match          5.4%; Score 40.8; DB 6; Length 713;
Best Local Similarity 49.1%; Pred. No. 0.18;
Matches 108; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 229 ATTAAGCGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 288
```

```
DB 301 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 360
QY 289 GTCATTTACAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 348
DB 361 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 420
QY 349 TTAGAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 408
DB 421 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 480
QY 409 TCCATTAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 448
DB 481 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 520

RESULT 13
US-10-363-345B-20053/c
; Sequence 20053, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 20053
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 20053
US-10-363-345B-20053

Query Match          5.3%; Score 40.6; DB 6; Length 786;
Best Local Similarity 49.8%; Pred. No. 0.21;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 227 TCATTAAAGCGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 286
DB 672 TTAATAAAGCGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 613
QY 287 AGGTCATTTACAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 346
DB 612 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 553
QY 347 TCTTAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 406
DB 552 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 493
QY 407 TATCCATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 433
DB 492 AATAAAGGAGGAGTCTTATACAAAGAGATTTTATGCGATGTGTGCGATAAATTAAAG 466

RESULT 14
US-10-363-345B-20054
; Sequence 20054, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363,345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 20054
; LENGTH: 786
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 20054
US-10-363-345B-20054

Query Match          5.3%; Score 40.6; DB 6; Length 786;
Best Local Similarity 49.8%; Pred. No. 0.21; Mismatches 0; Gaps 0;
Matches 103; Conservative 0; Indels 104;

QY 227 TCATTAAAGCGGAGGAGTCTTATACAAAGAGATTTTATCGATGTGTGCGATATAATTAA 286
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TTAATAACCGGAAATAAATAACCAAAAAACGATCGAAACGAAACGATAAAAAAT 174

QY 287 AGGTCAATTACAAAGAAACTGAAACGACTTTAATTGAACAAAAATGCTTTCTAAAA 346
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 AAAAAAATAAAAAAATAACGAAAAAATAAAAAACGCAAAAAAATCGCTACTACGAA 234

QY 347 TCTTAGAAGAGTTTGGAGAAATGATGATGAAGAGTGAAGAGATGCGGATGAAT 406
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 AAAATAAAAAAATAAAAAAATAACGAAAAAATAAAAACTAAAAAATAAAAAACGATAAA 294

QY 407 TATCCATAAAAAACACGACAAATTAA 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 AAAAAAATAAAAAAATAAAAAAATAAAAAA 321

RESULT 15
US-10-363-345B-31443/c
; Sequence 31443, Application US/10363345B
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: 82010
; CURRENT APPLICATION NUMBER: US/10/363.345B
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40723
; SEQ ID NO 31443
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 31443
US-10-363-345B-31443

Query Match          5.3%; Score 40.6; DB 6; Length 822;
Best Local Similarity 49.8%; Pred. No. 0.21;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 227 TCATTAAAGCGGAGGAGTCTTATACAAAGAGATTTTATCGATGTGTGCGATATAATTAA 286
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 TTAATAAACCGGAAATAAATAACCAAAAAACGATCGAAACGAAACGATAAAAAAT 740

QY 287 AGGTCAATTACAAAGAAACTGAAACGACTTTAATTGAACAAAAATGCTTTCTAAAA 346
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 AAAAAAATAAAAAAATAACGAAAAAATAAAAAACGCAAAAAAATCGCTACTACGAA 680

QY 347 TCTTAGAAGAGTTTGGAGAAATGATGATGAAGAGTGAAGAGATGCGGATGAAT 406
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 AAAATAAAAAAATAAAAAAATAACGAAAAAATAAAAACTAAAAAATAAAAAACGATAAA 620

QY 407 TATCCATAAAAAACACGACAAATTAA 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 AAAAAAATAAAAAAATAAAAAAATAAAAAA 593
```

Search completed: August 3, 2006, 10:58:47
Job time : 184 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2006, 12:03:35 ; Search time 49 Seconds
(without alignments)

451.944 Million cell updates/sec

Title: US-09-732-091-4

Perfect score: 1279

Sequence: 1 MAYKYDRDLFLKQLESSDL.....LRLKTOQANGKKSQIESI 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep:*
- 2: /EMC Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep:*
- 3: /EMC Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep:*
- 4: /EMC Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep:*
- 5: /EMC Celerra_SIDS3/ptodata/2/iaa/PCTUS_COMB.pep:*
- 6: /EMC Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep:*
- 7: /EMC Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	28.5	273	2	US-09-252-991A-29841
2	346.5	27.1	321	2	US-09-252-991A-29840
3	247	19.3	258	2	US-09-489-039A-8852
4	100	7.8	739	2	US-09-543-681A-6437
5	99	7.7	222	2	US-08-944-604-16
6	99	7.7	274	2	US-09-134-001C-5279
7	98	7.7	386	2	US-09-248-796A-17815
8	96	7.5	173	1	US-08-658-639-12
9	96	7.5	173	2	US-08-944-604-12
10	95	7.4	2125	2	US-09-919-172-29
11	95	7.4	2704	2	US-09-538-092-1260
12	93	7.3	381	2	US-09-710-279-3056
13	91.5	7.2	1086	2	US-09-543-681A-7696
14	90	7.0	216	2	US-09-710-279-2624
15	89.5	7.0	281	2	US-09-198-452A-749
16	89.5	7.0	281	2	US-09-438-185A-707
17	89.5	7.0	543	2	US-09-328-352-5845
18	89.5	7.0	546	2	US-09-345-236B-98
19	89.5	7.0	546	2	US-09-345-236B-121
20	89.5	7.0	560	2	US-09-446-301A-50
21	89.5	7.0	1402	2	US-09-248-796A-14503
22	88.5	6.9	552	2	US-09-446-301A-4
23	88.5	6.9	552	2	US-09-099-932-4
24	88.5	6.9	552	2	US-10-392-970-4
25	88	6.9	488	2	US-09-489-039A-13363
26	87.5	6.8	732	1	US-08-533-669A-18

27	87.5	6.8	732	2	US-09-307-143-4	Sequence 4, Appli
28	87.5	6.8	732	2	US-09-183-861-18	Sequence 18, Appl
29	87.5	6.8	732	2	US-09-022-765-18	Sequence 18, Appl
30	87.5	6.8	732	2	US-09-551-974A-18	Sequence 18, Appl
31	87.5	6.8	732	2	US-09-565-501A-18	Sequence 18, Appl
32	87.5	6.8	732	2	US-09-639-206A-18	Sequence 18, Appl
33	87.5	6.8	732	2	US-09-874-923-18	Sequence 18, Appl
34	87.5	6.8	732	2	US-08-798-841-18	Sequence 18, Appl
35	87.5	6.8	752	2	US-09-949-016-7993	Sequence 7993, Ap
36	87.5	6.8	854	2	US-10-094-749-2559	Sequence 2559, Ap
37	87	6.8	280	2	US-09-830-230A-596	Sequence 596, App
38	87	6.8	302	2	US-09-830-230A-595	Sequence 595, App
39	87	6.8	491	2	US-09-029-267-2	Sequence 2, Appli
40	87	6.8	434	2	US-09-248-796A-22920	Sequence 22920, A
41	86.5	6.8	331	2	US-09-134-001C-3626	Sequence 3626, Ap
42	86.5	6.8	420	2	US-09-248-796A-15323	Sequence 15323, A
43	86.5	6.8	1042	2	US-09-792-024-106	Sequence 106, App
44	86	6.7	1164	2	US-09-538-092-399	Sequence 399, App
45	85	6.6	183	2	US-09-270-767-33130	Sequence 33130, A

ALIGNMENTS

RESULT 1

US-09-252-991A-29841

; Sequence 29841, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 29841

; LENGTH: 273

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-29841

Query Match 28.5%; Score 364; DB 2; Length 273;

Best Local Similarity 35.4%; Pred. No. 1.9e-30;

Matches 85; Conservative 49; Mismatches 88; Indels 18; Gaps 4;

QY 10 EFLKQLESSDLLDFVLVFGKDGKRNHKLSTSSIEYKRGDDYAKYAEIAEELQYVG 69

DB 28 QLLERSVNDLLEPLVEYIL-----KARTESLSKQVDFKWHPEHRYASAILDELRLFG 81

QY 70 SNSFSFIKGEVLYKEILCDVCKLVNKNKTTTLIEQNMLSKILERSLEEMDDDEV 129

DB 82 GNSFANLWRRKSGPSYTEVVRDVGKLVKGVSGMELIELEAMVQSILRQALEKSSGEDR 141

QY 130 KEMCDEL-----SIKNTDNLNRQALSAAATLTFKMGFKSYQLAVIVANAVAKTILGRG 183

DB 142 RELEETLRAGLDKTKWTALLNGSALSG--LVVPAVARMILYRTSTVIVNSMAQQLLGHG 199

QY 184 LSLA-----GNQVLTRTSLFTGPVGMITGVWTAIDTAGPAYRVVTIPACIVVATLRUKTQ 239

DB 200 LRSVVAGGTGAGGRAVAALAGVGVWAGVWTAIDTAGPAYRVVTIPCVLHIALMLRKAR 259

RESULT 2

US-09-252-991A-29840

; Sequence 29840, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; CURRENT APPLICATION NUMBER: US/09/252.991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 29840
;; LENGTH: 321
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29840

Query Match 27.1%; Score 346.5; DB 2; Length 321;
Best Local Similarity 34.3%; Pred. No. 1.8e-28;
Matches 82; Conservative 42; Mismatches 112; Indels 3; Gaps 3;
QY 1 MAYKYDRDL-EFLKQLESLLDLFEVLVFGKDGKRNHKLTSSEYKRGHDDYAKYAE 59
DB 53 MAIHADLADLVLSASSDDIRLLIDVITDNGRISLSSVCRQLSAAKEGVVGEFERG 112
QY 60 RIAEELQYGSNSFASPIK-GEGLVYKEILCDVCDKLKVNKKTTETTLIEQNMLSKILE 118
DB 113 MVABELMRFGNSLMNLFRCGSGVPYKELSDVASHVGVSKESTGDCARMEMAIITKVE 172
QY 119 RSLEEMDDEEVKEMCDLSIKNTDNLNRQALSATLTLFRMGGPKSYQLAVIVANAVAKT 178
DB 173 QSIGRMEEDKATTFESIGTSYRSGMGPPVAL-AALIASLSASGWTSYGLAAMVASATMSS 231
QY 179 ILRGSLAGNQLVTRTLSPFTGPGVMIITGVTWTAIDAGPAYRVTTIPACIVATILK 237
DB 232 LVGRGVALAGGATIGRGLSLVTPGVGMAIAGIWTAFDLASPAYRVTLPCVIQIGHMRQK 290

RESULT 3
US-09-489-039A-8852
; Sequence 8852, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8852
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8852

Query Match 19.3%; Score 247; DB 2; Length 258;
Best Local Similarity 28.2%; Pred. No. 5.7e-18;
Matches 70; Conservative 48; Mismatches 100; Indels 30; Gaps 6;
QY 3 YKYDRDLFEFLKQLESLLDLFEVLVFGKDGKRNHKLTSSEYKRGHDDYAKYAE 59
DB 14 YLEDDDLAFLPECSAEHLFAFTRILTHGNGKPRLSSTLLRNETHFLAMEGHPERYRRNQ 73
QY 60 RIAEELQYGSNSFASPIKGEGLVYKEILCDVCDKLKVNKKTTETTLIEQNMLSKILER 119
DB 74 LIAGELQHFQGDSTANTLRHGRKPYRAILLDVCKRLKAKVDKQLSTPQIEQOLLAHFLQH 133
QY 120 SLEEMDDEEVKEMCDLSIKN--TDNL-----NRQALSATLTLFRMGGPKSYQLAVIV 171
DB 134 SWNKINAEQKQAFLAAVECRSHELDSLMAHLRHRKLGSEGVTLILL-----DERLTAIL 186

QY 172 ANAVAKTILGRGLSLAGNQVLTRTLSPFTGPGVMIITGVTWTAIDAGPAYRVTTIPACIV 231
DB 187 RTHAAVSVIGHGVLRGAG-----LGGPLGAALNSVKA---VSGSAYRVTTIPAVLHI 234
QY 232 ATLRLKTQ 239
DB 235 ACLRQMLQ 242
RESULT 4
US-09-543-681A-6437
; Sequence 6437, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6437
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6437

Query Match 7.8%; Score 100; DB 2; Length 739;
Best Local Similarity 25.6%; Pred. No. 0.15;
Matches 54; Conservative 42; Mismatches 65; Indels 50; Gaps 10;
QY 6 DRDLFEFLKQLESLLDLFEVLVFGKDGKRNHKLTSSEYKRGHDDYAKYAE 65
DB 517 DKVSDFLVSLVNSSLDGLKEL---GLD--KKIIEEMKNSIQ-----DKLKWATKILVT 565
QY 66 QYQGSNSFASPIKGEGLVYKEILCDVCDK-----LKVNKKTTETTL--IEQNMLS 114
DB 566 VLFVAATALSFVIGPAM---KQISDAVNKISNQRIQLKVLNDGLEAVLGMKIDIII 622
QY 115 KILERSLEEMDDEEVKEMCDLSIKNTDNLNRQAL-----SAATLTLFRMGGPKSYQLA 168
DB 623 KALEALEKID----KQLAKEISKASIMLNKTVVASKLTNSAATVTNVIYGS----- 671
QY 169 VIVANAVAKTILGRGLSLAGNQVLTRTLSPFL 199
DB 672 -VIASKIIQ-----SIAGSKKLTAVLDII 694

RESULT 5
US-08-944-604-16
; Sequence 16, Application US/08944604
; Patent No. 6218131
; GENERAL INFORMATION:
; APPLICANT: KEESSE, SUSAN
; APPLICANT: OEAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: MATERIALS AND METHODS FOR DETECTION OF
; TITLE OF INVENTION: BREAST CANCER
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,604
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, THOMAS C
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: MTP-021 (8395/24)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-944-604-16

Query Match
Best Local Similarity 21.4%; Score 99; DB 2; Length 222;
Matches 39; Conservative 48; Mismatches 83; Indels 12; Gaps 5;

QY 6 DRDLFLKQLESDDLDFEVLVFGK-DGEKRNKELTSSIEYKRGHDDYAKYAEIABE 64
Db 30 DRERQKLETQEKIADIKKMAKQGMQDAVRIMAKDLVTRTRYVRKFLMLRANIQAVALK 89
QY 65 LQYGS-NSPASFYKG----EGVLYKEILCDVCDKLVNKNKTTTTLIEQNMLSKILER 119
Db 90 IQTLKSNNSMAQMGVTKAMGTWNRQLKLPQIQIMMEFERQAEIMDKKEEMNDADD 149
QY 120 SLEEMDDEE----VKEMCELSIKNTDNLNRQALSAAATLTLFKMGGFKSYQLAVIVANA 174
Db 150 AWGDEDEEEDAVVSQVDELGLSLTDELSNLPSTGGSLV-AAGGKKAEEAASALADA 208
QY 175 VA 176
Db 209 DA 210

RESULT 6
US-09-134-001C-5279
; Sequence 5279, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5279
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-5279

Query Match
Best Local Similarity 23.8%; Score 99; DB 2; Length 274;
Matches 62; Conservative 44; Mismatches 88; Indels 66; Gaps 14;

QY 3 YKQDRDLFLKQLESDDLDFEVLVFGKDGKRNKELTSS-----IE 46
Db 4 YNNYNDL-----FLILKFGDEMKNQTIETISREALIKDVMMIARILLE 47
QY 47 YKRGHDDYAKYAEIABE LQYGSNSPASFYKGEGLVLYKEILCDVCDKLVNKNKTT-ET 105
Db 48 SGAEGTRVEDTMARIATKLGYPESNFVNTVIEFLVNLNEAYPRL-----YRIKTRDT 100

```

```

QY 106 TLIEQNMLSKILERSLE-EMDDEEVKEMCDELSI-KNTDNLNRQALS---ATLTLFRMG 160
Db 101 NLIKISOANEISRQITNGTWTLBEEKYOLEEIVYAKRDSSLPFKGIAAAIATSFLYQQ 160
QY 161 GPKSYQLAVIVANAVAKTILGRGLSLAGNQLVTRL--SELTGPVGMIIIGVWTAIDIA 218
Db 161 G----RLVDIITAVLAGTI---GYLVV--BILDRKLAHQFIPFIFIGSLVIGI---ISVIG 208
QY 219 PAYRVTPIP---ACIVVATL 234
Db 209 HAF---VPSGDLATIIIAAV 225

RESULT 7
US-09-248-796A-17815
; Sequence 17815, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17815
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-17815

Query Match
Best Local Similarity 7.7%; Score 98; DB 2; Length 386;
Matches 52; Conservative 35; Mismatches 84; Indels 46; Gaps 10;

```

```

QY 12 LKQL--ESSDLLD-----LFEVLVFGKDGKRNKELTSSIEYKRGHDDYAKYAEIABE 64
Db 61 LKELTAEQLEILLDGPVDILFRTL-----QEVDPITIAKKFHPQDHRKL--RRALE 107
QY 65 LQYGSNSPASFYKGE--GVLYKEILCDVCDKLVNKNKTTTTLIEQNMLSKILERSLE 122
Db 108 IFY-----TKGEKASEIYHEQKLELDSSSLKYNTLFFWVYCDPEILNDRDKRVD 158
QY 123 EMDD---EEVKEMCELSIKNTDNLNRQALSAAATLTLFKMGGFKSYQLAVIVANAVAKT 178
Db 159 KMWENGAIEIKEMVDFYKSKQEQNL-----TCTSGIWOVIGPKFF-LLWLEDNQDITK 211
QY 179 ILGRGLSLAGNQLVTRLTSLTGPVGMIIIGVWTAID 215
Db 212 LFEHGI----ERMKIRTOYARYQVQVKKWSLLTELE 244

RESULT 8
US-08-658-639-12
; Sequence 12, Application US/08658639
; Patent No. 5914238
; GENERAL INFORMATION:
; APPLICANT: KEESSE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: MATERIALS AND METHODS FOR DETECTION OF
; TITLE OF INVENTION: BREAST CANCER
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA

```



```

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurPatSeqFormatter Version 0.9
; SEQ ID NO 1260
; LENGTH: 2704
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q03001
; US-09-538-092-1260

```

Query Match	7.4%;	Score 95;	DB 2;	Length 2704;
Best Local Similarity	25.4%;	Pred. No. 3.7;		
Matches	46;	Conservative 35;	Mismatches 62;	Indels 38; Gaps 9;
QY	2	AYKYDRDLFLK--QLESSDLLDF-----EVLVFGKDGKEKRHEKLTSSIEYKRHGDY	55	
DB	1565	SFRDEKELERLQICQKSDHLKQFPEKSHQQLLQNIKAERKENDKIQRLNEELEKSN	1624	
QY	56	KYAEARIAEEL--QYGSNSPASTKIGSG---VLVKEILCDVCDKLKV-----	97	
DB	1625	EMLKQKVEELTRQNNETKLMQRIQABSENIVLEKQTIQORCEALKIQADGFKDQLRSTN	1684	
QY	98	-NYNKKTETTLIQNMLSKI--LERSILEMDD--EYKEMCDELSI-----KNTDNLNR	146	
DB	1685	EHHLQTKT---EQDFORKICLEEDLAKSNLVSEFKQCDQONIIIQNTKKEVRNLNA	1741	
QY	147	Q	147	
DB	1742	E	1742	

```

RESULT 12
US-09-710-279-3056
; Sequence 3056, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3056
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-3056

```

```

Query Match      7.3%; Score 93; DB 2; Length 381;
Best Local Similarity 24.0%; Pred. No. 0.31;
Matches 53; Conservative 30; Mismatches 74; Indels 64; Gaps 10;

QY  4 KYDRDLFLKQLRSSLDLLFVLVFGKGEKGRHNEKLTSSIBYKRGHGDYAK--YAERI 61
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db  19 KVDHNEIYLKTTTKDKSFLDMDTI-----SNLYSKFICYGFGEFYQQDLLNLNYQIER 71

QY  62 AELQVYGSN--SPASFTKGEGVLYKEILCDVCKLKVYNKKTETTLI-----EQNMLSKIL 117

```

Db	72	QNVLQEFNSDPNIQNF	: : : :	DEML	: : : : :	NKLQKVSLLISASEESTKKIV	112
Qy	118	ERSLEEDMDDEVKEMCDLSIKNTDNLNRQALSAAATLFLPWGGFKSYQLAVIVAN-AVA	: : : : :				176
Db	113	DHFVEELYSEEPKQKIN	: : : : :			---TGVKLVVDYKIGGLEPTQLIVIAARPSVG	157
Qy	177	KT	-----ILGRG-----	-----LSLAGNQVLTRTLFLTG			201
Db	158	KTFPALNMLNLIASOGYKTFSPFSLETTGVSVLKRLMSAETG	: : : : :				198

```

RESULT 13
US-09-543-681A-7696
; Sequence 7696, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELAT
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7696
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7696

```

Query Match	7.2%	Score 91.5;	DB 2;	Length 1086;
Best local Similarity	19.5%;	Pred. No. 2.2;		
Matches 43;	Conservative	38;	Mismatches 73;	Indels 67; Gaps 9;
Qy	3	YKDRDLFLKQLQESSDLLDLFEVLVFGDKGKRRHNEKLTSSIEYKRHGDDYAKYAERIA	62	
Db	786	FAYDRGKE-IEQMQ-----FETTLIGK--SRAEQEKINA-----	816	
Qy	63	EELQYYGSNSFASFIKGEVLYKEILCDVCDKULKVYNKKTETTTLEQNMLSKILERSLE	122	
Db	817	-----LRQIDVLYQQASVDLGEKELVNLQRNVLT--KQIEEELRKRKEAM	860	
Qy	123	EMD-----DREVKEMCDLSIKNTDNLNRQALSAAATLTLPKMGFKSYQLAVI	170	
Db	861	KGDPMAGLKQGLSDPFSESAMDVENVNTTNALNNMSDALADALFALTCKSGFKDPANAVI	920	
Qy	171	--VANAVAKTILGRGSLAGNQVLTRTLTSLFTGPVGVMIITG	209	
Db	921	SDITRWVMKMLIFKATKAEAG-----QAMGF---DMGMSKSG	953	

```

RESULT 14
US-09-710-279-2624
; Sequence 2624, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU34805H
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2624
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic

```

; OTHER INFORMATION: amino acid sequence
US-09-710-279-2624

Query Match 7.0%; Score 90; DB 2; Length 216;
Best Local Similarity 27.8%; Pred. No. 0.28;
Matches 52; Conservative 35; Mismatches 66; Indels 34; Gaps 12;
QY 60 RIARELQYGSNSPASKGEGVLYKEILCDVCKLVNKNKT-ETTLIEQNMLSKILE 118
DB 3 RIATKLGYPESNFVTWTFVLFVHNAYPLP-----YRIKTRDTNLKISOANEISR 55
QY 119 RSLE-EMDDEBEVKMCDLSI-KNTDNLNRQALSA--ATLTFMGGFKSYQLAVIVAN 173
DB 56 QITNGTWTLEAKYQLEEIYVAKEDSSLPPFGIAAAIATSFYIQGG---RLVDIITA 111
QY 174 AVAKTILGRGLSLAGNOVLRTLL--SFLTGPVGMHITGVWTAIDIAGPAYRVTP----A 227
DB 112 VLACTI---GYLVV--EILDRLKHAQTPFPGSLVIGI---ISVIGHAF---VPSGDLA 160
QY 228 CIVVATL 234
DB 161 TIIIAAV 167

RESULT 15

US-09-198-452A-749
; Sequence 749, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198.452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 749
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-749

Query Match 7.0%; Score 89.5; DB 2; Length 281;
Best Local Similarity 22.0%; Pred. No. 0.46;
Matches 56; Conservative 49; Mismatches 103; Indels 47; Gaps 11;
QY 22 DLFEVLVFGDKGKRHNEKLTSS-----IEYKRHGGDYAKYAEIAE-----ELQYY 68
DB 30 DSRDVKFVLEGGKQTRQETKTSSKGNTRTESKRFADDEKRVDDDEIAEVGSKBEESQEF 89
QY 69 --GSNSFA--SFIKGEGVLYKEILCDVCKLVNKNKTETTLIEQNMLSKILERSLEEM 124
DB 90 CLAEAFAGNLIIDIAAGSAEAVVEVA-PIAVS---SIDTQWIENIILSTVESWVISEI 145
QY 125 DDEVKEMCDLSIKNTDNLNRQALSAATLTFMGGFKSYQLAVIVANAKTILGRGL 184
DB 146 NGEQLVEL-----VLDASSVPEAFVGANLTLVQSG-----QDLSVKFSFVDATQMAEA 196
QY 185 SLAGN--QVLTRTLSTLTG-----PVGWIIICVWTAIDIAGPAYRVTPACIVVATL 234
DB 197 DLVTNNPQSLSLVSALKGHQTLTKFSGVGNLLVQLPKIEVQTPLEH-----MIASTI 249
QY 235 RLKTOOANGDKKSLQ 249
DB 250 RHREKQORDQONQK 264

Search completed: August 3, 2006, 12:04:50
Job time : 50 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2006, 12:15:20 ; Search time 176 Seconds
(without alignments)

665.871 Million cell updates/sec

Title: US-09-732-091-4

Perfect score: 1279

Sequence: 1 MAYKYDRDLEFLKQLESSDL.....LRLKTOQANGDKKSLQIESI 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1279	100.0	253	3	US-09-732-091-4
2	1279	100.0	253	4	US-10-433-970-4
3	1279	100.0	265	3	US-09-732-091-44
4	1279	100.0	265	4	US-10-433-970-44
5	1270	99.3	253	4	US-10-335-977-9162
6	1270	99.3	253	4	US-10-335-977-9163
7	1270	99.3	256	4	US-10-335-977-9164
8	1264	98.8	253	4	US-10-433-970-48
9	1264	98.8	265	4	US-10-433-970-45
10	722	56.5	248	4	US-10-335-977-7699
11	716	56.0	237	4	US-10-335-977-7698
12	472.5	36.9	155	3	US-09-882-227-414
13	252	19.7	49	3	US-09-732-091-20
14	252	19.7	49	4	US-10-433-970-20
15	198	15.5	38	3	US-09-732-091-17
16	198	15.5	38	4	US-10-433-970-17
17	194	15.2	41	3	US-09-732-091-19
18	194	15.2	41	4	US-10-433-970-19
19	152	11.9	30	3	US-09-732-091-18
20	152	11.9	30	4	US-10-433-970-18
21	150	11.7	30	3	US-09-732-091-16
22	150	11.7	30	4	US-10-433-970-16
23	99	7.7	274	4	US-10-724-972A-5261
24	98.5	7.7	1009	4	US-10-282-122A-43832
25	97.5	7.6	815	4	US-10-437-963-191043
26	97	7.6	916	4	US-10-282-122A-76490
27	96.5	7.5	373	4	US-10-363-493-42

ALIGNMENTS

RESULT 1

US-09-732-091-4

; Sequence 4, Application US/09732091

; Patent No. US20020107368A1

; GENERAL INFORMATION:

; APPLICANT: Tian, Jing-Hui

; APPLICANT: Walker, Richard I.

; APPLICANT: Jackson, W. James

; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses

; TITLE OF INVENTION: thereof

; FILE REFERENCE: 7969-088

; CURRENT APPLICATION NUMBER: US/09/732,091

; CURRENT FILING DATE: 2000-12-07

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Helicobacter sp.

US-09-732-091-4

Query Match	100.0%;	Score 1279;	DB 3;	Length 253;
Best Local Similarity	100.0%;	Pred. No. 3.7e-108;		
Matches 253;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAYKYDRDLEFLKQLESSDL	1	MAYKYDRDLEFLKQLESSDL
Db	1	MAYKYDRDLEFLKQLESSDL	1	MAYKYDRDLEFLKQLESSDL
QY	61	IAEELQYGSNSFASFIKGEVLYKEILCDVCDKLVNKNKTTETTLIEQNMLSKILERS	120	IAEELQYGSNSFASFIKGEVLYKEILCDVCDKLVNKNKTTETTLIEQNMLSKILERS
Db	61	IAEELQYGSNSFASFIKGEVLYKEILCDVCDKLVNKNKTTETTLIEQNMLSKILERS	120	IAEELQYGSNSFASFIKGEVLYKEILCDVCDKLVNKNKTTETTLIEQNMLSKILERS
QY	121	LEEMDEEVEKMCDELISKNTDNLNRQALSAATLTFLKMGGFKSYQLAVIVANAVAKTIL	180	LEEMDEEVEKMCDELISKNTDNLNRQALSAATLTFLKMGGFKSYQLAVIVANAVAKTIL
Db	121	LEEMDEEVEKMCDELISKNTDNLNRQALSAATLTFLKMGGFKSYQLAVIVANAVAKTIL	180	LEEMDEEVEKMCDELISKNTDNLNRQALSAATLTFLKMGGFKSYQLAVIVANAVAKTIL
QY	181	GRGLSLAGNQVLTRTLFLTLTGPVGMITGTGTAIDAGPAYRVITIPACIVVATLRLKTOQ	240	GRGLSLAGNQVLTRTLFLTLTGPVGMITGTGTAIDAGPAYRVITIPACIVVATLRLKTOQ
Db	181	GRGLSLAGNQVLTRTLFLTLTGPVGMITGTGTAIDAGPAYRVITIPACIVVATLRLKTOQ	240	GRGLSLAGNQVLTRTLFLTLTGPVGMITGTGTAIDAGPAYRVITIPACIVVATLRLKTOQ
QY	241	ANGDKKSLQIESI 253		
Db	241	ANGDKKSLQIESI 253		

RESULT 2

US-10-433-970-4

; Sequence 4, Application US/10433970

; Publication No. US20040138415A1

```
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-10-433-970-4
Query Match 100.0%; Score 1279; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.7e-108;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAYKYDRDLFLKQLESSDLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 60
DB 1 MAYKYDRDLFLKQLESSDLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 60
QY 61 IAEELQYGSNSFASFIKGEGVLYKEILCDVCDKLVKNYKKTETTLIEQNMLSKILERS 120
DB 61 IAEELQYGSNSFASFIKGEGVLYKEILCDVCDKLVKNYKKTETTLIEQNMLSKILERS 120
QY 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
DB 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDAGPAYRVTIIPACIVVATLRLKTOQ 240
DB 181 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDAGPAYRVTIIPACIVVATLRLKTOQ 240
QY 193 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDAGPAYRVTIIPACIVVATLRLKTOQ 252
DB 193 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDAGPAYRVTIIPACIVVATLRLKTOQ 252
QY 241 ANGDKKSLQIESI 253
DB 241 ANGDKKSLQIESI 253
; RESULT 4
US-10-433-970-44
; Sequence 44, Application US/10433970
; Publication No. US20040138415A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-10-433-970-44
Query Match 100.0%; Score 1279; DB 4; Length 265;
Best Local Similarity 100.0%; Pred. No. 4e-108;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAYKYDRDLFLKQLESSDLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 60
DB 13 MAYKYDRDLFLKQLESSDLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 72
QY 61 IAEELQYGSNSFASFIKGEGVLYKEILCDVCDKLVKNYKKTETTLIEQNMLSKILERS 120
DB 73 IAEELQYGSNSFASFIKGEGVLYKEILCDVCDKLVKNYKKTETTLIEQNMLSKILERS 132
QY 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
DB 133 LEEMDDEEVKEMCDLSIKNTDNLNRQALSATLTLFKMGGFKSYQLAVIVANAVAKTIL 192
QY 181 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDAGPAYRVTIIPACIVVATLRLKTOQ 240
DB 193 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDAGPAYRVTIIPACIVVATLRLKTOQ 252
QY 241 ANGDKKSLQIESI 253
DB 253 ANGDKKSLQIESI 265
; RESULT 5
US-10-335-977-9162
; Sequence 9162, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO HELICOBACTER PYLORI FOR
```

DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 9162:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...253
SEQUENCE DESCRIPTION: SEQ ID NO: 9162:
US-10-335-977-9162

Query Match 99.3%; Score 1270; DB 4; Length 253;
Best Local Similarity 99.2%; Pred. No. 2.5e-107;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLFLKQLESDDLDFVLVFGKDGKGRHNEKLTSSIEYKRHGGDDYAKYAE 60
DB 1 MAYKYDRDLFLKQLESDDLDFVLVFGKDGKGRHNEKLTSSIEYKRHGGDDYAKYAE 60
QY 61 IAEELQYGSNSPASFIFKGEVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
DB 61 IAEELQYGSNSPASFIFKGEVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
QY 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
DB 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNQVLTSLTSLTGPVGMIIIGVWTAIDTAGPAYRVITIPACIVVATLRLKTTQ 240
DB 181 GRGLSLAGNQVLTSLTSLTGPVGMIIIGVWTAIDTAGPAYRVITIPACIVVATLRLKTTQ 240
QY 241 ANGDKKSLQIESI 253
DB 241 ANEDKKSQIESV 253

RESULT 6
US-10-335-977-9163
; Sequence 9163, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 9163:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...253
SEQUENCE DESCRIPTION: SEQ ID NO: 9163:
US-10-335-977-9163

Query Match 99.3%; Score 1270; DB 4; Length 253;
Best Local Similarity 99.2%; Pred. No. 2.5e-107;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLFLKQLESDDLDFVLVFGKDGKGRHNEKLTSSIEYKRHGGDDYAKYAE 60
DB 1 MAYKYDRDLFLKQLESDDLDFVLVFGKDGKGRHNEKLTSSIEYKRHGGDDYAKYAE 60
QY 61 IAEELQYGSNSPASFIFKGEVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
DB 61 IAEELQYGSNSPASFIFKGEVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
QY 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
DB 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNQVLTSLTSLTGPVGMIIIGVWTAIDTAGPAYRVITIPACIVVATLRLKTTQ 240
DB 181 GRGLSLAGNQVLTSLTSLTGPVGMIIIGVWTAIDTAGPAYRVITIPACIVVATLRLKTTQ 240
QY 241 ANGDKKSLQIESI 253
DB 241 ANEDKKSQIESV 253

RESULT 7
US-10-335-977-9164
; Sequence 9164, Application US/10335977
; Publication No. US20040052799A1

GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 9164:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...256
SEQUENCE DESCRIPTION: SEQ ID NO: 9164:
US-10-335-977-9164

Query Match 99.3%; Score 1270; DB 4; Length 256;
Best Local Similarity 99.2%; Pred. No. 2.5e-107;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKRRHNEKLTSSIEYKRHGGDDYAKYAE 60
DB 4 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKRRHNEKLTSSIEYKRHGGDDYAKYAE 63
QY 61 IAEELQYYSNSFASFIKGEGLVYKEILCDVCDKLVNKKTTTTLIEQNMLSKILERS 120
DB 64 IAEELQYYSNSFASFIKGEGLVYKEILCDVCDKLVNKKTTTTLIEQNMLSKILERS 123
QY 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLFLFMGGPKSYQLAVIVANAVAKTIL 180
DB 124 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLFLFMGGPKSYQLAVIVANAVAKTIL 183
QY 181 GRGLSLAGNOVLRTLSTLFTGPVGMIIITGVMTAIDAGPAYRVTTIPACIVVATLRLKTTQ 240
DB 184 GRGLSLAGNOVLRTLSTLFTGPVGMIIITGVMTAIDAGPAYRVTTIPACIVVATLRLKTTQ 243
QY 241 ANGDKKSLQIESI 253
DB 244 ANEDKKSLQIESV 256
RESULT 8
US-10-433-970-48

Sequence 48, Application US/10433970
Publication No. US20040138415A1
GENERAL INFORMATION:
APPLICANT: Tian, Jing-Hui
APPLICANT: Walker, Richard
APPLICANT: Jackson, James
TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
FILE REFERENCE: 7969-091-999
CURRENT APPLICATION NUMBER: US/10/433,970
CURRENT FILING DATE: 2003-06-06
PRIOR APPLICATION NUMBER: 09/732,091
PRIOR FILING DATE: 2000-12-07
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 48
LENGTH: 253
TYPE: PRT
ORGANISM: Helicobacter sp.
US-10-433-970-48

Query Match 98.8%; Score 1264; DB 4; Length 253;
Best Local Similarity 98.8%; Pred. No. 8.7e-107;
Matches 250; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKRRHNEKLTSSIEYKRHGGDDYAKYAE 60
DB 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKRRHNEKLTSSIEYKRHGGDDYAKYAE 60
QY 61 IAEELQYYSNSFASFIKGEGLVYKEILCDVCDKLVNKKTTTTLIEQNMLSKILERS 120
DB 61 IAEELQYYSNSFASFIKGEGLVYKEILCDVCDKLVNKKTTTTLIEQNMLSKILERS 120
QY 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLFLFMGGPKSYQLAVIVANAVAKTIL 180
DB 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLFLFMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNOVLRTLSTLFTGPVGMIIITGVMTAIDAGPAYRVTTIPACIVVATLRLKTTQ 240
DB 181 GRGLSLAGNOVLRTLSTLFTGPVGMIIITGVMTAIDAGPAYRVTTIPACIVVATLRLKTTQ 240
QY 241 ANGDKKSLQIESI 253
DB 241 ANEDKKSLQIESI 253

RESULT 9
US-10-433-970-46
Sequence 46, Application US/10433970
Publication No. US20040138415A1
GENERAL INFORMATION:
APPLICANT: Tian, Jing-Hui
APPLICANT: Walker, Richard
APPLICANT: Jackson, James
TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
FILE REFERENCE: 7969-091-999
CURRENT APPLICATION NUMBER: US/10/433,970
CURRENT FILING DATE: 2003-06-06
PRIOR APPLICATION NUMBER: 09/732,091
PRIOR FILING DATE: 2000-12-07
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 265
TYPE: PRT
ORGANISM: Helicobacter sp.
US-10-433-970-46

Query Match 98.8%; Score 1264; DB 4; Length 265;
Best Local Similarity 98.8%; Pred. No. 9.3e-107;
Matches 250; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLFLKQLBSSDLLDFVLVFGKDGKRNHKLTSSEYKRGDDYAKYAE 60
DB 13 MAYKYDRDLFLKQLBSSDLLDFVLVFGKDGKRNHKLTSSEYKRGDDYAKYAE 72
QY 61 IAEELQYGSNFPASPIKGEVLYKEILCDVCDKLVNKKYKTTTLLIQNMLSKILERS 120
DB 73 IAEELQYGSNFPASPIKGEVLYKEILCDVCDKLVNKKYKTTTLLIQNMLSKILERS 132
QY 121 LBEEMDEEYKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
DB 133 LBEEMDEEYKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 192
QY 181 GRGLSLAGNQLTRTSLFTGPGVMIITGVWTAIDAGPAYRTIPACIVVATLRKTOQ 240
DB 193 GRGLSLAGNQLTRTSLFTGPGVMIITGVWTAIDAGPAYRTIPACIVVATLRKTOQ 252
QY 241 ANGDKKSQIESI 253
DB 253 ANEDKKSQIESI 265

RESULT 10
US-10-335-977-7699
; Sequence 7699, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 7699:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...248
; SEQUENCE DESCRIPTION: SEQ ID NO: 7699:
US-10-335-977-7699

Query Match 56.5%; Score 722; DB 4; Length 248;
Best Local Similarity 60.3%; Pred. No. 2.3e-57;

Matches 141; Conservative 36; Mismatches 53; Indels 4; Gaps 2;
QY 8 DLEFLKQLBSSDLLDFVLVFGKDGKRNHKLTSSEYKRGDDYAKYAEELQY 67
DB 11 DLEFLKQLBSSDLLDFVLVFGKDGKRNHKLTSSEYKRGDDYAKYAEELQY 70
QY 68 YGSNFPASPIKGEVLYKEILCDVCDKLVNKKYKTTTLLIQNMLSKILERSLEEMDDE 127
DB 71 YGSNFPASPIKGEVLYKEILCDVCDKLVNKKYKTTTLLIQNMLSKILERSLEEMDDE 130
QY 128 EYKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTILGRGL 184
DB 131 EYKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTILGRGL 190
QY 185 -SLAGNQLTRTSLFTGPGVMIITGVWTAIDAGPAYRTIPACIVVATLRK 237
DB 191 -SLAGNQLTRTSLFTGPGVMIITGVWTAIDAGPAYRTIPACIVVATLRK 244

RESULT 11
US-10-335-977-7698
; Sequence 7698, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 7698:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...237
; SEQUENCE DESCRIPTION: SEQ ID NO: 7698:
US-10-335-977-7698

Query Match 56.0%; Score 716; DB 4; Length 237;
Best Local Similarity 60.1%; Pred. No. 7.7e-57;
Matches 140; Conservative 36; Mismatches 53; Indels 4; Gaps 2;

```
QY 9 LEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRRHGGDDYAKYAEIRIAEELQYY 68
Db 1 LEFLKRLSSDDLKDLFDALVYDEGTLRMEELTSLTEYQRYGHDYAKYPRRIAEELQRY 60
QY 69 GSNFSASFPIKGBGVLYKEILCDVCDKJLVNKNKTTETTLTLEQNMLSKILERSLEEMDDE 128
Db 61 GSNFSANFFRDEGVLYKEILCDACDHLDDINYNERSATSLIEQNMLSKLLKDSLEKMSGRE 120
QY 129 VKEMCDLSIKNTDNL---NRQALSATLTLFRMGKPKSYOLAVIVANAVAKTTLGRCL- 184
Db 121 IKELCDGLGPNFDKVTGENKQVLIASVLTFLKAGGSHSYALAVADAMVRQTLGHGLS 180
QY 185 SLAGNQVLTTLSLTPGVGWIITGVWTAIDIAGPAYRVVTIPACIVVATLRK 237
Db 181 SVGKVALKKTLDILAGPIGWITGALVSINLAGPAYRVVTPACVIVATLRKK 233

RESULT 12
US-09-882-227-414
; Sequence 414, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Anas
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 414
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-882-227-414

Query Match 36.9%; Score 472.5; DB 3; Length 155;
Best Local Similarity 63.3%; Pred. No. 7e-35;
Matches 93; Conservative 22; Mismatches 29; Indels 3; Gaps 1;

QY 38 NEKLTSSIEYKRRHGGDDYAKYAEIRIAEELQYYGSNSFASFIKGBGVLYKEILCDVCDKLV 97
Db 2 NEDLTNSTEYKRYGHDYAKYPRRIAEELQRYGGSNSFANFFRDEGVLYKEILCDACDHLKV 61
QY 98 NYNKKTETTLIEQNMLSKILERSLEEMDDEEVKEMCDLSIKNTDNL---NRQALSATL 154
Db 62 NYNEESATSLIEQNMLSKLLKDSLEKMSRREIKELCNELGWNTIDKVTGENKQVLIASLT 121
QY 155 TLFKMGKPKSYOLAVIVANAVAKTILG 181
Db 122 TLFKAGGSHSYALAVSADAMVRQTLG 148

RESULT 13
US-09-732-091-20
; Sequence 20, Application US/09732091
; Patent No. US20020107368A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-09-732-091-20

Query Match 15.5%; Score 198; DB 3; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-09-732-091-20

Query Match 19.7%; Score 252; DB 3; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.9e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 VGMIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQANGDKKSLOIE 251
Db 1 VGMIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQANGDKKSLOIE 49

RESULT 14
US-10-433-970-20
; Sequence 20, Application US/10433970
; Publication No. US20040138415A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-10-433-970-20

Query Match 19.7%; Score 252; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.9e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 VGMIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQANGDKKSLOIE 251
Db 1 VGMIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQANGDKKSLOIE 49

RESULT 15
US-09-732-091-17
; Sequence 17, Application US/09732091
; Patent No. US20020107368A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-09-732-091-17

Query Match 15.5%; Score 198; DB 3; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	53 DVAKYAERIAEELQYVGSNFSASIKGEGVLKYEILCD 90
Db	1 DVAKYAERIAEELQYVGSNFSASIKGEGVLKYEILCD 38

Search completed: August 3, 2006, 12:18:40
Job time : 177 secs

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2006, 12:15:50 ; Search time 31 Seconds
(without alignments)
546.105 Million cell updates/sec

Title: US-09-732-091-4
Perfect score: 1279
Sequence: 1 MAYKYDRDLFLKQLLESSDL.....LRLKTOQANGDKKSLQIESI 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 236815 seqs, 66914042 residues

Total number of hits satisfying chosen parameters: 236815

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	98	7.7	413	6	US-10-471-571A-1526	Sequence 1526, Ap
2	96.5	7.5	282	6	US-10-449-902-33147	Sequence 33147, A
3	94	7.3	576	6	US-10-531-357A-2	Sequence 2, Appl
4	91.5	7.2	187	6	US-10-953-349-4873	Sequence 4873, Ap
5	91.5	7.2	187	7	US-11-056-355B-43591	Sequence 43591, A
6	91.5	7.2	187	7	US-11-056-355B-98508	Sequence 98508, A
7	91.5	7.2	187	7	US-11-056-355B-109747	Sequence 109747, A
8	91.5	7.2	241	6	US-10-953-349-4872	Sequence 4872, Ap
9	91.5	7.2	241	7	US-11-056-355B-43590	Sequence 43590, A
10	91.5	7.2	241	7	US-11-056-355B-98507	Sequence 98507, A
11	91.5	7.2	241	7	US-11-056-355B-109746	Sequence 109746, A
12	91.5	7.2	257	7	US-11-056-355B-98506	Sequence 98506, A
13	91.5	7.2	257	7	US-11-056-355B-109745	Sequence 109745, A
14	91.5	7.2	261	6	US-10-953-349-4871	Sequence 4871, Ap
15	91.5	7.2	261	7	US-11-056-355B-43589	Sequence 43589, A
16	87.5	6.8	462	7	US-11-056-355B-83110	Sequence 83110, A
17	87.5	6.8	489	7	US-11-056-355B-83109	Sequence 83109, A
18	87.5	6.8	732	7	US-11-105-233-171	Sequence 171, App
19	87	6.8	247	6	US-10-471-571A-972	Sequence 972, App
20	87	6.8	402	6	US-10-471-571A-2786	Sequence 2786, Ap
21	87	6.8	1182	7	US-11-056-355B-48406	Sequence 48406, A
22	87	6.8	1200	7	US-11-056-355B-48405	Sequence 48405, A
23	87	6.8	1294	7	US-11-056-355B-48404	Sequence 48404, A
24	85.5	6.7	457	7	US-11-330-403-3995	Sequence 3995, Ap
25	84.5	6.6	454	7	US-11-056-355B-80275	Sequence 80275, A

ALIGNMENTS

RESULT 1

US-10-471-571A-1526 483 7 US-11-056-355B-80274 Sequence 80274, A
; Sequence 1526, Application US/10471571A 286 7 US-11-246-405-23 Sequence 23, Appl
; Publication No. US20060115490A1 286 7 US-11-365-556-354 Sequence 354, App
; GENERAL INFORMATION: 371 7 US-11-330-403-9557 Sequence 9557, Ap
; APPLICANT: CHIRON SPA 443 7 US-11-330-403-8304 Sequence 8304, Ap
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS 597 7 US-11-056-355B-48918 Sequence 48918, A
; FILE REFERENCE: P02692WO 628 7 US-11-056-355B-48917 Sequence 48917, A
; CURRENT APPLICATION NUMBER: US/10/471,571A 633 7 US-11-056-355B-48916 Sequence 48916, A
; PRIOR FILING DATE: 2003-09-12 407 7 US-11-330-403-14767 Sequence 14767, A
; PRIOR APPLICATION NUMBER: GB-0107661.1 435 6 US-10-953-349-2242 Sequence 2242, Ap
; PRIOR FILING DATE: 2001-03-27 435 7 US-11-056-355B-38279 Sequence 38279, A
; NUMBER OF SEQ ID NOS: 5642 435 7 US-11-056-355B-106125 Sequence 106125, A
; SOFTWARE: Seqwin99, version 1.03 435 7 US-11-056-355B-117364 Sequence 117364, A
; SEQ ID NO 1526 435 7 US-11-056-355B-2241 Sequence 2241, Ap
; LENGTH: 413 500 6 US-10-953-349-2241 Sequence 2241, Ap
; TYPE: PRT 500 7 US-11-056-355B-38278 Sequence 38278, A
; ORGANISM: Staphylococcus aureus 500 7 US-11-056-355B-106124 Sequence 106124, A
; FEATURE: 500 7 US-11-056-355B-117363 Sequence 117363, A
; NAME/KEY: MISC_FEATURE 506 7 US-11-056-355B-106123 Sequence 106123, A
; LOCATION: (1)..(413) 506 7 US-11-056-355B-117362 Sequence 117362, A
; OTHER INFORMATION: replicative DNA helicase 510 6 US-10-953-349-2240 Sequence 2240, Ap
US-10-471-571A-1526

Query Match 7.7%; Score 98; DB 6; Length 413;
Best Local Similarity 24.0%; Pred. No. 2;
Matches 53; Conservative 26; Mismatches 78; Indels 64; Gaps 8;
QY 4 KYDRDLFLKQLLESSDLLD-----LFEVLVFGKDGKGRHNEKLTSTSIYKRRHGDDYAKY 57
DB 50 KIDHQEIYUKATKDEFLDADTTIKLYNSDFICYGFERYQOELLESYQLN-----KA 102
QY 58 ABRIAEELQYGSNSFASPIKGEGVLYKEILCDVCDKLKVNYNKTTETTLIEONMLSKIL 117
DB 103 NELVTEFKQPTNQNFNNLI-----DELKDLKTTITNKEDGT-----KKFV 143
QY 118 ERSLEEMDEEVKEMCDLSIKNTDNLNRQALSAAITLTFKMGGFKSYQLAVIVAN-AVA 176
DB 144 BEFVEEL-----YDSPFKQIKTGKLYMDYKIGGLEPSQLIVIAARPSVG 188
QY 177 KT-----TLGRG-----LSLAGNQVLRTRLSFLTG 201
DB 189 KTGFALNMLNIAQNGYKTSFFSLETTGTSLVKRLMLSTTG 229

RESULT 2

US-10-449-902-33147
; Sequence 33147, Application US/10449902

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2006, 12:04:10 ; Search time 604 Seconds
(without alignments)
639.603 Million cell updates/sec

Title: US-09-732-091-4
Perfect score: 1279
Sequence: 1 MAYKYDRDLFLKLESSL.....LRKLTQANGDKKSLQIESI 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCTUS COMB.pcp.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066 COMB.pcp.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073 COMB.pcp.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/paa/US074 COMB.pcp.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075 COMB.pcp.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076 COMB.pcp.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077 COMB.pcp.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078 COMB.pcp.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079 COMB.pcp.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080 COMB.pcp.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081 COMB.pcp.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082 COMB.pcp.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083 COMB.pcp.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084 COMB.pcp.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085 COMB.pcp.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086 COMB.pcp.*
- 17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087 COMB.pcp.*
- 18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088 COMB.pcp.*
- 19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089 COMB.pcp.*
- 20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090 COMB.pcp.*
- 21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091 COMB.pcp.*
- 22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092 COMB.pcp.*
- 23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093 COMB.pcp.*
- 24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094 COMB.pcp.*
- 25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095 COMB.pcp.*
- 26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096 COMB.pcp.*
- 27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097 COMB.pcp.*
- 28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098 COMB.pcp.*
- 29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099 COMB.pcp.*
- 30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100 COMB.pcp.*
- 31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101 COMB.pcp.*
- 32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102 COMB.pcp.*
- 33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103 COMB.pcp.*
- 34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104 COMB.pcp.*
- 35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105 COMB.pcp.*
- 36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106 COMB.pcp.*
- 37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107 COMB.pcp.*
- 38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108 COMB.pcp.*
- 39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109 COMB.pcp.*
- 40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110 COMB.pcp.*
- 41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111 COMB.pcp.*
- 42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112 COMB.pcp.*
- 43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113 COMB.pcp.*

- 44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114 COMB.pcp.*
- 45: /EMC_Celerra_SIDS3/ptodata/2/paa/US600 COMB.pcp.*
- 46: /EMC_Celerra_SIDS3/ptodata/2/paa/US601 COMB.pcp.*
- 47: /EMC_Celerra_SIDS3/ptodata/2/paa/US602 COMB.pcp.*
- 48: /EMC_Celerra_SIDS3/ptodata/2/paa/US603 COMB.pcp.*
- 49: /EMC_Celerra_SIDS3/ptodata/2/paa/US604 COMB.pcp.*
- 50: /EMC_Celerra_SIDS3/ptodata/2/paa/US605 COMB.pcp.*
- 51: /EMC_Celerra_SIDS3/ptodata/2/paa/US606 COMB.pcp.*
- 52: /EMC_Celerra_SIDS3/ptodata/2/paa/US607 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1279	100.0	253	1	PCT-US01-48392-4
2	1279	100.0	253	27	US-09-732-091-4
3	1279	100.0	253	34	US-10-433-970-4
4	1279	100.0	265	1	PCT-US01-48392-44
5	1279	100.0	265	27	US-09-732-091-44
6	1279	100.0	265	34	US-10-433-970-44
7	1270	99.3	253	1	PCT-US96-18542-205
8	1270	99.3	253	14	US-08-487-032A-785
9	1270	99.3	253	15	US-08-561-469A-785
10	1270	99.3	253	17	US-08-761-184-870
11	1270	99.3	253	17	US-08-761-184-1471
12	1270	99.3	253	18	US-08-821-931-870
13	1270	99.3	253	18	US-08-821-931-1471
14	1270	99.3	253	19	US-08-993-002A-9162
15	1270	99.3	253	19	US-08-993-002A-9163
16	1270	99.3	253	33	US-10-335-977-9162
17	1270	99.3	253	33	US-10-335-977-9163
18	1270	99.3	256	17	US-08-761-184-1215
19	1270	99.3	256	18	US-08-821-931-1215
20	1270	99.3	256	19	US-08-993-002A-9164
21	1270	99.3	256	33	US-10-335-977-9164
22	1264	98.8	253	1	PCT-US01-48392-48
23	1264	98.8	253	34	US-10-433-970-48
24	1264	98.8	265	1	PCT-US01-48392-46
25	1264	98.8	265	34	US-10-433-970-46
26	722	56.5	248	19	US-08-993-002A-7699
27	722	56.5	248	33	US-10-335-977-7699
28	716	56.0	237	19	US-08-993-002A-7698
29	716	56.0	237	33	US-10-335-977-7698
30	615	48.1	209	14	US-08-487-032C-831
31	472.5	36.9	155	1	PCT-US98-06371-1148
32	472.5	36.9	155	19	US-08-902-615A-414
33	472.5	36.9	155	28	US-09-882-227-414
34	364	28.5	273	33	US-10-366-683-29841
35	364	28.5	273	34	US-10-419-128-29841
36	346.5	27.1	321	33	US-10-366-683-29840
37	346.5	27.1	321	34	US-10-419-128-29840
38	252	19.7	49	1	PCT-US01-48392-20
39	252	19.7	49	27	US-09-732-091-20
40	252	19.7	49	34	US-10-433-970-20
41	247	19.3	258	34	US-10-446-203-8852
42	198	15.5	38	1	PCT-US01-48392-17
43	198	15.5	38	27	US-09-732-091-17
44	198	15.5	38	34	US-10-433-970-17
45	194	15.2	41	1	PCT-US01-48392-19

ALIGNMENTS

RESULT 1
PCT-US01-48392-4
; Sequence 4, Application PC/TUS0148392
; GENERAL INFORMATION:

; APPLICANT: AntexBiologicals, Inc.
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; FILE OF INVENTION: thereof
; FILE REFERENCE: 7969-091-228
; CURRENT APPLICATION NUMBER: PCT/US01/48392
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Helicobacter sp.
PCT-US01-48392-4

Query Match 100.0%; Score 1279; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 4.8e-118;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
DB 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
QY 61 IAEELQYGSNSFASPIKGEVLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
DB 61 IAEELQYGSNSFASPIKGEVLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
QY 121 LEEMDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
DB 121 LEEMDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIITGVWTAIDIAIAGPAYRVITPACIVVATLRLKTOQ 240
DB 181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIITGVWTAIDIAIAGPAYRVITPACIVVATLRLKTOQ 240
QY 241 ANGDKKSLQIESI 253
DB 241 ANGDKKSLQIESI 253

RESULT 2

US-09-732-091-4

; Sequence 4, Application US/09732091
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; FILE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-09-732-091-4

Query Match 100.0%; Score 1279; DB 27; Length 253;
Best Local Similarity 100.0%; Pred. No. 4.8e-118;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
DB 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
QY 61 IAEELQYGSNSFASPIKGEVLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
DB 61 IAEELQYGSNSFASPIKGEVLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120

QY 121 LEEMDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
DB 121 LEEMDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIITGVWTAIDIAIAGPAYRVITPACIVVATLRLKTOQ 240
DB 181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIITGVWTAIDIAIAGPAYRVITPACIVVATLRLKTOQ 240
QY 241 ANGDKKSLQIESI 253
DB 241 ANGDKKSLQIESI 253

RESULT 3

US-10-433-970-4

; Sequence 4, Application US/10433970
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard
; APPLICANT: Jackson, James
; TITLE OF INVENTION: HELICOBACTER PROTEINS, GENE SEQUENCES AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: 7969-091-999
; CURRENT APPLICATION NUMBER: US/10/433,970
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-10-433-970-4

Query Match 100.0%; Score 1279; DB 34; Length 253;
Best Local Similarity 100.0%; Pred. No. 4.8e-118;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
DB 1 MAYKYDRDLFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
QY 61 IAEELQYGSNSFASPIKGEVLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
DB 61 IAEELQYGSNSFASPIKGEVLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
QY 121 LEEMDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
DB 121 LEEMDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIITGVWTAIDIAIAGPAYRVITPACIVVATLRLKTOQ 240
DB 181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIITGVWTAIDIAIAGPAYRVITPACIVVATLRLKTOQ 240
QY 241 ANGDKKSLQIESI 253
DB 241 ANGDKKSLQIESI 253

RESULT 4

PCT-US01-48392-44

; Sequence 44, Application PC/TUS0148392
; GENERAL INFORMATION:
; APPLICANT: AntexBiologicals, Inc.
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; FILE OF INVENTION: thereof
; FILE REFERENCE: 7969-091-228
; CURRENT APPLICATION NUMBER: PCT/US01/48392
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/732,091
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 48

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Helicobacter sp.
PCT-US01-48392-44

Query Match      100.0%; Score 1279; DB 1; Length 265;
Best Local Similarity 100.0%; Pred. No. 5.2e-118;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAYKYDRDLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 60
Db      13 MAYKYDRDLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 72

QY      61 IAELOYGNSGNSFASFIKGEGLVYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
Db      73 IAELOYGNSGNSFASFIKGEGLVYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 132

QY      121 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFMGGFKSYQLAVIVANAVAKTIL 180
Db      133 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFMGGFKSYQLAVIVANAVAKTIL 192

QY      181 GRGLSLAGNOVLRTLSFLTGPVGWIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQ 240
Db      193 GRGLSLAGNOVLRTLSFLTGPVGWIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQ 252

QY      241 ANGDKKSQIQESI 253
Db      253 ANGDKKSQIQESI 265

RESULT 5
US-09-732-091-44
; Sequence 44, Application US/09732091
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Helicobacter sp.
US-09-732-091-44

Query Match      100.0%; Score 1279; DB 27; Length 265;
Best Local Similarity 100.0%; Pred. No. 5.2e-118;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAYKYDRDLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 60
Db      13 MAYKYDRDLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRHGDDYAKYAE 72

QY      61 IAELOYGNSGNSFASFIKGEGLVYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
Db      73 IAELOYGNSGNSFASFIKGEGLVYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 132

QY      121 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFMGGFKSYQLAVIVANAVAKTIL 180
Db      133 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAAATLTLFMGGFKSYQLAVIVANAVAKTIL 192

QY      181 GRGLSLAGNOVLRTLSFLTGPVGWIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQ 240
Db      193 GRGLSLAGNOVLRTLSFLTGPVGWIIITGVWTAIDIAGPAYRVVTIPACIVVATLRKLTQQ 252

QY      241 ANGDKKSQIQESI 253

```

```
; APPLICATION NUMBER: PCT/US96/18542
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/561,469
; FILING DATE: 17-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001CPPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 205:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORGANISM: Helicobacter pylori
; NAME/KEY: misc feature
; LOCATION: 1...253
PCT-US96-18542-205

Query Match 99.3%; Score 1270; DB 1; Length 253;
Best Local Similarity 99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLFLKQLESSLDDLFLVLFVKDGEKRRNEKLTSSIEYKRGHGGDDYAKYAE 60
Db 1 MAYKYDRDLFLKQLESSLDDLFLVLFVKDGEKRRNEKLTSSIEYKRGHGGDDYAKYAE 60
QY 61 IAEELQYIGNSFASFIKGEGLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
Db 61 IAEELQYIGNSFASFIKGEGLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
QY 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
Db 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNOVLTRTLSPFTGPGVGIITGWTATIDAGPAYRVTIPACIIVATLRLKTOQ 240
Db 181 GRGLSLAGNOVLTRTLSPFTGPGVGIITGWTATIDAGPAYRVTIPACIIVATLRLKTOQ 240
QY 241 ANGDKKSLQIESI 253
Db 241 ANEDKKSQIESV 253

RESULT 8
US-08-487-032A-785
; Sequence 785, Application US/08487032A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 880
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/561,469A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; APPLICATION NUMBER: US/08/487,032A
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 785:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORGANISM: Helicobacter pylori
; NAME/KEY: misc feature
; LOCATION: 1...253
US-08-487-032A-785

Query Match 99.3%; Score 1270; DB 14; Length 253;
Best Local Similarity 99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLFLKQLESSLDDLFLVLFVKDGEKRRNEKLTSSIEYKRGHGGDDYAKYAE 60
Db 1 MAYKYDRDLFLKQLESSLDDLFLVLFVKDGEKRRNEKLTSSIEYKRGHGGDDYAKYAE 60
QY 61 IAEELQYIGNSFASFIKGEGLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
Db 61 IAEELQYIGNSFASFIKGEGLYKEILCDVCDKLVNKKTTTLLIEQNMLSKILERS 120
QY 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
Db 121 LEEMDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
QY 181 GRGLSLAGNOVLTRTLSPFTGPGVGIITGWTATIDAGPAYRVTIPACIIVATLRLKTOQ 240
Db 181 GRGLSLAGNOVLTRTLSPFTGPGVGIITGWTATIDAGPAYRVTIPACIIVATLRLKTOQ 240
QY 241 ANGDKKSLQIESI 253
Db 241 ANEDKKSQIESV 253

RESULT 9
US-08-561-469A-785
; Sequence 785, Application US/08561469A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 994
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/561,469A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
```

NAME: Mandragoras, Amy E.	US 08/630,405
REGISTRATION NUMBER: 36,207	FILING DATE: 01-APR-1996
REFERENCE/DOCKET NUMBER: GTN-001CP	PRIOR APPLICATION DATA:
TELECOMMUNICATION INFORMATION:	TELEPHONE: (617) 227-7400
TELEPHONE: (617) 227-7400	TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 785:	SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids	TYPE: amino acid
TOPOLOGY: linear	MOLECULE TYPE: protein
HYPOTHETICAL: YES	ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori	NAME/KEY: misc feature
US-08-561-469A-785	LOCATION: 1...253
Query Match	99.3%; Score 1270; DB 15; Length 253;
Best Local Similarity	99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative	1; Mismatches 1; Indels 0; Gaps 0;
Qy	1 MAYKYDRDLFLKQLLESSDLLDLFEVLVFGKDGKRRHNEKLTSSIEYKRHGGDDYAKYAE 60
Db	1 MAYKYDRDLFLKQLLESSDLLDLFEVLVFGKDGKRRHNEKLTSSIEYKRHGGDDYAKYAE 60
Qy	61 IAELOYYGNSFPASPIKGEVLYKEILCDVCDKLVNKNKTTTLLIEQNMLSKILERS 120
Db	61 IAELOYYGNSFPASPIKGEVLYKEILCDVCDKLVNKNKTTTLLIEQNMLSKILERS 120
Qy	121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
Db	121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAAATLTLFKMGGPKSYQLAVIVANAVAKTIL 180
Qy	181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIIGVWTAIDIAGPAYRVITPACIVVATRLKTKQ 240
Db	181 GRGLSLAGNQLVTRTSLFSLTGPVGMIIIGVWTAIDIAGPAYRVITPACIVVATRLKTKQ 240
Qy	241 ANGDKKSLOIESI 253
Db	241 ANEDKKSLOIESV 253
RESULT 11	
US-08-761-184-1471	Sequence 1471, Application US/08761184
GENERAL INFORMATION:	
APPLICANT: DOUGLAS SMITH ET AL	
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES	
TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI AND	
TITLE OF INVENTION: THERAPEUTIC USES THEREOF	
NUMBER OF SEQUENCES: 1810	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: LAHIVE & COCKFIELD	
STREET: 60 State Street, Suite 510	
CITY: Boston	
STATE: Massachusetts	
COUNTRY: USA	
ZIP: 02109-1875	
COMPUTER READABLE FORM:	
MEDIUM TYPE: CD/ROM ISO9660	
COMPUTER:	
OPERATING SYSTEM:	
SOFTWARE:	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/08/761,184	
FILING DATE:	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: US 08/487,032	
FILING DATE: 07-JUN-1995	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: US 08/561,469	
FILING DATE: 17-NOV-1995	
PRIOR APPLICATION DATA:	

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,184
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/561,469
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,405
; FILING DATE: 01-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/660,742
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/736,791
; FILING DATE: 25-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/739,150
; FILING DATE: 28-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001CP9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1471:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..253
US-08-761-184-1471

Query Match          99.3%; Score 1270; DB 17; Length 253;
Best Local Similarity 99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
DB 1 MAYKYDRDLLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60

QY 61 TAEELQYYSNSFASFIKGEVLYKEILCDVCDKLVKNVYKKTETTLIEQNMLSKILERS 120
DB 61 TAEELQYYSNSFASFIKGEVLYKEILCDVCDKLVKNVYKKTETTLIEQNMLSKILERS 120

QY 121 LEEMDEEVKEMDELSIKNTDNLNQAALSAATLTLFKMGKPSYQLAVIVANAKTLL 180
DB 121 LEEMDEEVKEMDELSIKNTDNLNQAALSAATLTLFKMGKPSYQLAVIVANAKTLL 180

QY 181 GRGLSLAGNQLVTRLTSLFTGPGVWITGTWTAIDAGPAYRTVPACIVWATRLKTOQ 240
DB 181 GRGLSLAGNQLVTRLTSLFTGPGVWITGTWTAIDAGPAYRTVPACIVWATRLKTOQ 240

QY 241 ANGDKKSLQIESI 253
DB 241 ANEDKKSLQIESV 253

US-08-821-931-870
; Sequence 870, Application US/08821931
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH ET AL
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; NUMBER OF SEQUENCES: 1810
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,931
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/561,469
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,405
; FILING DATE: 01-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/660,742
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/736,791
; FILING DATE: 25-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/739,150
; FILING DATE: 28-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001CP9CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 870:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..253
US-08-821-931-870

Query Match          99.3%; Score 1270; DB 18; Length 253;
Best Local Similarity 99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60
DB 1 MAYKYDRDLLEFLKQLESSDLLDLFEVLVFGKDGKGRHNEKLTSSIEYKRGHGGDYAKYAE 60

RESULT 12
```

Qy 61 IAEELQYGSNSPASFIKGEVLYKEILCDVCDKLVKNYNKKTETTTLIEQNMLSKILERS 120
Db 61 IAEELQYGSNSPASFIKGEVLYKEILCDVCDKLVKNYNKKTETTTLIEQNMLSKILERS 120
Qy 121 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
Db 121 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
Qy 181 GRGLSLAGNOVLTRTSLFSLTGPVGMIIITGVWTAIDIAGPAYRVITIPACIVVATLRLKTOQ 240
Db 181 GRGLSLAGNOVLTRTSLFSLTGPVGMIIITGVWTAIDIAGPAYRVITIPACIVVATLRLKTOQ 240
Qy 241 ANGDKKSLQIESI 253
Db 241 ANEDKSLQIESV 253

RESULT 13

US-08-821-931-1471
; Sequence 1471, Application US/08821931
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH ET AL
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; NUMBER OF SEQUENCES: 1810
; CORRESPONDENCE ADDRESS: 1810
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,931
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,032
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/561,469
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,405
; FILING DATE: 01-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/660,742
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/736,791
; FILING DATE: 25-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/739,150
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/761,184
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001CP9CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1471:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1...253
US-08-821-931-1471
Query Match 99.3%; Score 1270; DB 18; Length 253;
Best Local Similarity 99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MAYKYDRDLFLKQLSSDLLDLFEVLVFGKDEKRNKLTSSIEYKRGHGGDYAKYAE 60
Db 1 MAYKYDRDLFLKQLSSDLLDLFEVLVFGKDEKRNKLTSSIEYKRGHGGDYAKYAE 60
Qy 61 IAEELQYGSNSPASFIKGEVLYKEILCDVCDKLVKNYNKKTETTTLIEQNMLSKILERS 120
Db 61 IAEELQYGSNSPASFIKGEVLYKEILCDVCDKLVKNYNKKTETTTLIEQNMLSKILERS 120
Qy 121 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
Db 121 LEEMDDEEVKEMCDELSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
Qy 181 GRGLSLAGNOVLTRTSLFSLTGPVGMIIITGVWTAIDIAGPAYRVITIPACIVVATLRLKTOQ 240
Db 181 GRGLSLAGNOVLTRTSLFSLTGPVGMIIITGVWTAIDIAGPAYRVITIPACIVVATLRLKTOQ 240
Qy 241 ANGDKKSLQIESI 253
Db 241 ANEDKSLQIESV 253

RESULT 14

US-08-993-002A-9162
; Sequence 9162, Application US/08993002A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS: 10031
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,002A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; INFORMATION FOR SEQ ID NO: 9162:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...253
US-08-993-002A-9162

Query Match      99.3%; Score 1270; DB 19; Length 253;
Best Local Similarity 99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLFLKQLESSDLDLFEVLVFGKGKRNHNEKLTSSIEYKRHGDDYAKYAE 60
   |||||
DB 1 MAYKYDRDLFLKQLESSDLDLFEVLVFGKGKRNHNEKLTSSIEYKRHGDDYAKYAE 60
   |||||

QY 61 IAEELQYYGSGNSFASFIKGEGVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
   |||||
DB 61 IAEELQYYGSGNSFASFIKGEGVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
   |||||

QY 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
   |||||
DB 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
   |||||

QY 181 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDTAGPAYRVTIIPACIVVATLRLKTOQ 240
   |||||
DB 181 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDTAGPAYRVTIIPACIVVATLRLKTOQ 240
   |||||

QY 241 ANGDKKSLQIESI 253
   |||||
DB 241 ANEDKKSQIESV 253
   |||||
```

RESULT 15

```
US-08-993-002A-9163
; Sequence 9163, Application US/08993002A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,002A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...253
US-08-993-002A-9163

Query Match      99.3%; Score 1270; DB 19; Length 253;
Best Local Similarity 99.2%; Pred. No. 3.8e-117;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYKYDRDLFLKQLESSDLDLFEVLVFGKGKRNHNEKLTSSIEYKRHGDDYAKYAE 60
   |||||
DB 1 MAYKYDRDLFLKQLESSDLDLFEVLVFGKGKRNHNEKLTSSIEYKRHGDDYAKYAE 60
   |||||

QY 61 IAEELQYYGSGNSFASFIKGEGVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
   |||||
DB 61 IAEELQYYGSGNSFASFIKGEGVLYKEILCDVCDKLKVNYNKKTETTLIEQNMLSKILERS 120
   |||||

QY 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
   |||||
DB 121 LEEMDDEEVKEMCDLSIKNTDNLNRQALSAATLTLFKMGGFKSYQLAVIVANAVAKTIL 180
   |||||

QY 181 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDTAGPAYRVTIIPACIVVATLRLKTOQ 240
   |||||
DB 181 GRGLSLAGNOVLTRTSLFTGPGVMIITGVWTAIDTAGPAYRVTIIPACIVVATLRLKTOQ 240
   |||||

QY 241 ANGDKKSLQIESI 253
   |||||
DB 241 ANEDKKSQIESV 253
   |||||
```

Search completed: August 3, 2006, 12:14:59
Job time : 605 secs

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	97.5	7.6	311	1	PCT-US06-18535-32605	Sequence 32605, A
2	97.5	7.6	311	7	US-11-431-855-32605	Sequence 32605, A
3	95	7.4	2649	7	US-11-188-417A-25	Sequence 25, Appl
4	92	7.2	309	7	US-11-434-137-9630	Sequence 9630, Ap
5	92	7.2	309	7	US-11-434-184-9630	Sequence 9630, Ap
6	92	7.2	309	7	US-11-434-203-9630	Sequence 9630, Ap
7	92	7.2	309	7	US-11-434-127-9630	Sequence 9630, Ap
8	92	7.2	309	7	US-11-434-199-9630	Sequence 9630, Ap
9	92	7.2	327	7	US-11-434-137-5500	Sequence 5500, Ap
10	92	7.2	327	7	US-11-434-184-5500	Sequence 5500, Ap
11	92	7.2	327	7	US-11-434-203-5500	Sequence 5500, Ap
12	92	7.2	327	7	US-11-434-127-5500	Sequence 5500, Ap
13	92	7.2	327	7	US-11-434-199-5500	Sequence 5500, Ap
14	91.5	7.2	241	7	US-11-479-226-1446	Sequence 1446, Ap
15	91.5	7.2	375	7	US-11-434-137-10134	Sequence 10134, A
16	91.5	7.2	375	7	US-11-434-184-10134	Sequence 10134, A
17	91.5	7.2	375	7	US-11-434-203-10134	Sequence 10134, A
18	91.5	7.2	375	7	US-11-434-127-10134	Sequence 10134, A
19	91.5	7.2	375	7	US-11-434-199-10134	Sequence 10134, A
20	91.5	7.2	390	7	US-11-434-137-4944	Sequence 4944, Ap
21	91.5	7.2	390	7	US-11-434-184-4944	Sequence 4944, Ap
22	91.5	7.2	390	7	US-11-434-203-4944	Sequence 4944, Ap
23	91.5	7.2	390	7	US-11-434-127-4944	Sequence 4944, Ap
24	91.5	7.2	390	7	US-11-434-199-4944	Sequence 4944, Ap
25	90	7.0	271	7	US-11-365-965-569	Sequence 569, App

```

RESULT 2
US-11-431-855-32605
; Sequence 32605, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32605
; LENGTH: 311
; TYPE: PRT
;

```

```

; ORGANISM: Campylobacter jejuni subsp. jejuni NCTC 11168
US-11-431-855-32605

Query Match          7.6%; Score 97.5; DB 7; Length 311;
Best Local Similarity 23.0%; Pred. No. 1.1;
Matches 37; Conservative 34; Mismatches 37; Indels 53; Gaps 8;

Qy 40 KLTSSIEYKHH-----GDDYAKYAERIAEELQYY---GNSNFASFIFKGEGLYKEIL 88
      : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 138 RIILTLGGKHGIIGLTIGKEVAKISKARGAEIYYSTSGANKNADFVHLE---LKDLL 194
      : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :

Qy 89 CDVCDKLKN--YNNKTETTL-----IEQMLSKILER----- 119
      || : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 195 -KTCDIISIAPLNKTKNLAFELKLLKDNAILNVGRGGIVNENDLAKIIDKRNIRV 253
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :

Qy 120 SLEEMDDEEYKWCDELSIKNTDNL-----NRCALSA 151
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 254 GLDVLEIEPMKNHPLLSIKNKENLIITPHVAMASKEALNA 294
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :

RESULT 3
US-11-188-417A-25
; Sequence 25, Application US/11188417A
; GENERAL INFORMATION:
; APPLICANT: CHAMPION, BRIAN ROBERT
; TITLE OF INVENTION: TREATMENT OF AUTOIMMUNE DISEASES USING AN ACTIVATOR FOR
; FILE REFERENCE: 674525-2022
; CURRENT APPLICATION NUMBER: US/11/188,417A
; CURRENT FILING DATE: 2005-07-25
; PRIOR APPLICATION NUMBER: PCT/GB04/00263
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: GB 0301519.5
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: GB 0301518.7
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: GB 0301515.3
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: GB 0301513.8
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: GB 0301512.0
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: GB 0301510.4
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: GB 0301524.5
; PRIOR FILING DATE: 2003-01-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 25
; LENGTH: 2649
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-188-417A-25

Query Match          7.4%; Score 95; DB 7; Length 2649;
Best Local Similarity 25.4%; Pred. No. 34;
Matches 46; Conservative 35; Mismatches 62; Indels 38; Gaps 9;

Qy 2 AYKYDRDLFLK--QLSSDDLDF-----EVLVFGKGEKRRHNEKLTSSIEYKHHGDDYA 55
      : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 1510 SFRDEKELERLOICQKSDHLKEQFEKSHEQLQNIKAENKNDKIQRLENELEKSNECA 1569
      : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :

Qy 56 KYAERIAEEL--QYVGSNSFASFYKGE--VLYKEILCDVCDKLKV----- 97
      : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 1570 EMLKQKVEELTRONNETKMMQRQAESENVLEKQTIQORCEALKIQADGFKQOLRSTN 1629
      : : : : | : : : : | : : : : | : : : : | : : : : | : : : :

Qy 98 -NNYKTKETTLIEQNMLSKI--LERSLEEMDD--EEVKEMDELSI-----KNTDNLNR 146
      : : : : | : : : : | : : : : | : : : : | : : : : | : : : :

```


QY 189 NOVLTRTSLTGPVGMIIITGVMTAI 214
Db 236 MRLVGAKNSYIRGP--FFPEGAWGI 259

RESULT 11
US-11-434-203-5500
; Sequence 5500, Application US/11434203
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,203
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 5500
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-11-434-203-5500

Query Match 7.2%; Score 92; DB 7; Length 327;
Best Local Similarity 21.8%; Pred. No. 3.6;
Matches 45; Conservative 35; Mismatches 60; Indels 66; Gaps 11;

QY 31 KDGKRNH-----EKLTSIEYKRGHDDYAKYAERIAEELQYGSNSFA 74
Db 98 KDKLNDPDYHKYDKIRISGVEKT-----YSSKAEQLKEVQKEVGSDDVID 146
QY 75 SFIKGEGVLYKEILCDVCDKLVNKNKTTTLLIEQNMLSKILERSLEEMDDEEVKEMCD 134
Db 147 D-----TYKDALLDV-----YVGTSSAKVSKS-VSEAIGR-IEGVD--YTKPEID 188
QY 135 ELSIKN-TDNLRQALSAATLTFKMGGPKSYQLAVIVA-----NAVAKTILGRGLSLAG 188
Db 189 STKLSNLTNI-----RIWFGGVALLIVLAIFLISNTIRMSRRTDIEI 235
QY 189 NOVLTRTSLTGPVGMIIITGVMTAI 214
Db 236 MRLVGAKNSYIRGP--FFPEGAWGI 259

RESULT 12
US-11-434-127-5500
; Sequence 5500, Application US/11434127
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,127
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 5500
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-11-434-127-5500

QY 31 KDGKRNH-----EKLTSIEYKRGHDDYAKYAERIAEELQYGSNSFA 74
Db 98 KDKLNDPDYHKYDKIRISGVEKT-----YSSKAEQLKEVQKEVGSDDVID 146
QY 75 SFIKGEGVLYKEILCDVCDKLVNKNKTTTLLIEQNMLSKILERSLEEMDDEEVKEMCD 134
Db 147 D-----TYKDALLDV-----YVGTSSAKVSKS-VSEAIGR-IEGVD--YTKPEID 188
QY 135 ELSIKN-TDNLRQALSAATLTFKMGGPKSYQLAVIVA-----NAVAKTILGRGLSLAG 188
Db 189 STKLSNLTNI-----RIWFGGVALLIVLAIFLISNTIRMSRRTDIEI 235
QY 189 NOVLTRTSLTGPVGMIIITGVMTAI 214
Db 236 MRLVGAKNSYIRGP--FFPEGAWGI 259

RESULT 13
US-11-434-199-5500
; Sequence 5500, Application US/11434199
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,199
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 5500
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-11-434-199-5500

Query Match 7.2%; Score 92; DB 7; Length 327;
Best Local Similarity 21.8%; Pred. No. 3.6;
Matches 45; Conservative 35; Mismatches 60; Indels 66; Gaps 11;

QY 31 KDGKRNH-----EKLTSIEYKRGHDDYAKYAERIAEELQYGSNSFA 74

```
Db 98 KDKLKNPDYHKYVDKIKRISGEKVT-----YSSKABQLAEVQKEYGSDVID 146
Qy 75 SFIKGBGVLYKEILCDVCDKLVNNYKKTETTLIEQNMLSKILERSLEEMDDEEVKEMCD 134
Db 147 D-----YTKDALLDV-----YVVGTSAAKVS-KVSEAIGR-IEGVD--YTKKEPID 188
Qy 135 ELSIKN-TDNLNRQALSATLTLPKMGGFKSVQLAVIVA-----NAVAKTILGRGLSLAG 188
Db 189 STKLSNLTDNI-----RINFGGVALIVLAIFLISNIRMSRRTDIEI 235
Qy 189 NOVLTRTSLFSLTGPVGIITGWTAI 214
Db 236 MRLVGAKNYSIRGP--PFPEGAWGCI 259
```

```
RESULT 14
US-11-479-226-1446
; Sequence 1446 Application US/11479226
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: REUBER, T. Lynne
; APPLICANT: RATCLIFFE, Oliver J
; APPLICANT: RUFF, Thomas G
; APPLICANT: THOMPSON-MIZE, Rebecca L
; APPLICANT: SUN, Jindong
; TITLE OF INVENTION: TRANSCRIPTION FACTORS FOR INCREASING YIELD
; FILE REFERENCE: MBI-0022-2CIP
; CURRENT FILING DATE: 2006-06-30
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/166,228
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/197,899
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: 60/227,439
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1446
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1030 Amino Acid Sequence
US-11-479-226-1446
```

```
Query Match 7.2%; Score 91.5; DB 7; Length 241;
Best Local Similarity 23.4%; Pred. No. 2.6;
Matches 33; Conservative 25; Mismatches 44; Indels 39; Gaps 5;

Qy 34 EKRHNEKLTSSIEYK-----RHGDDVAKYAEIRIAEELQYYGSNS 72
Db 100 KKTNDKSSSTSNKPKRPLTAPFPMSPDKTFKEHNGSLAKDAKI-----GGEK 152

Qy 73 FASFIKGBGVLYKEILCDVCDKLVNNYKKTETTLIEQNMLSKILERSLEEMDDEEVKEM 132
Db 153 WKSLETEE---KKVYLDKAAELKAEYKNSLESNDADDEEEDB--EKQSDDDVDDAEKQV 206

Qy 133 CD-----ELSIKNTDNLNRQA 148
Db 207 DDDDEVEKEVENTDDDKKEA 227
```

```
RESULT 15
US-11-434-137-10134
; Sequence 10134, Application US/11434137
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
```

```
; APPLICANT: Tettelin, Herve
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,137
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: Seqwin99, version 1.02
; SEQ ID NO 10134
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-11-434-137-10134

Query Match 7.2%; Score 91.5; DB 7; Length 375;
Best Local Similarity 25.7%; Pred. No. 4.7;
Matches 37; Conservative 33; Mismatches 61; Indels 13; Gaps 7;

Qy 74 ASFIKGBGVLYKEILCDVCDK-LKVNKKTTETTLIEQNMLSKILERSLEEMDDEEVKEM 132
Db 228 SSVFKGGGTIYLDVLPNISVKELEIHFQNEREBSL--KNALKILLPERLAEFYAEIDLPEK 285
Qy 133 CDELSIKNTDNLNRQALSATLTLPKMGGFKSV--QLAVIVANAVAKTILGR---GLSLA 187
Db 286 VKQSVKDLLEMLTIQKKLPILVTGKMSLAKSPVTKGGVDLKEINPKTLESKKVAGLHFA 345
Qy 188 GN--QVLTRTSLF-LTGPV--GWI 206
Db 346 GEVLIDINAHTGGFNITSALCTGVV 369

Search completed: August 3, 2006, 12:15:38
Job time : 34 secs
```

72798 8-1078 199339 M9

STIC-Biotech/ChemLib

From: Portner, Ginny
Sent: Monday, August 21, 2006 5:35 PM
To: STIC-Biotech/ChemLib
Subject: 10/387,977

please interference search SEQ ID NO 1 and 2. thanks!!

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

8/22/2006



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 199339

TO: Ginny Portner
Location: rem/3B02/3C18
Art Unit: 1645
August 31, 2006

Case Serial Number: 10/387977

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 17:49:50 ; Search time 215.372 Seconds
(without alignments)
44.581 Million cell updates/sec

Title: US-10-387-977-1

Perfect score: 118

Sequence: 1 FNGGISLANYTGHGSETAWGT 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.8.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*
10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	118	100.0	21	2	AAW83070	Peptide f
2	118	100.0	493	2	AAW83125	Peptide f
3	118	100.0	507	2	AAW83124	PtIIIR45
4	118	100.0	1706	2	AAW24786	PtIIIR50
5	118	100.0	1706	9	ADZ10438	P. gingiv
6	118	100.0	1706	9	ADZ10403	P. gingiv
7	118	100.0	1706	9	ADZ10427	P. gingiv
8	118	100.0	1706	9	ADZ10456	P. gingiv
9	118	100.0	1706	9	AEBS91466	Microbial
10	114	96.6	435	4	AAW49555	Gingipain
11	114	96.6	492	7	ABW02694	Porphyrom
12	114	96.6	492	9	ADV16838	Human cys
13	114	96.6	507	7	ABW02695	Porphyrom
14	114	96.6	737	2	AAW70186	Arg-gingi
15	114	96.6	737	2	AAW34846	Arg-gingi
16	114	96.6	737	3	AAW67395	Arg-gingi
17	114	96.6	737	4	AAU08937	P. gingiv
18	114	96.6	991	2	AAW77313	Porphyrom
19	114	96.6	1477	7	ABW02693	Porphyrom
20	114	96.6	1687	2	AAW96033	P. gingiv
21	114	96.6	1687	2	AAW69495	Haemagglu
22	114	96.6	1704	2	AAW70188	Arg-gingi
23	114	96.6	1704	2	AAW34843	Arg-gingi

ALIGNMENTS

RESULT 1
AAW83070

ID AAW83070 standard; peptide; 21 AA.

AC AAW83070;

DT 03-MAR-1999 (first entry)

DE Peptide from the 300kDa complex of cysteine proteinases and adhesins.

300 kD multiprotein complex; cysteine proteinase; adhesin;
virulence factor; immune response; P. gingivalis infection;
periodontal disease; passive immunisation; subgingival plaque.

OS Synthetic.

OS Porphyromonas gingivalis.

PN WO9849192-A1.

PD 05-NOV-1998.

PF 30-APR-1998; 98WO-AU000311.

PR 30-APR-1997; 97AU-00006528.

PA (UYME) UNIV MELBOURNE.
(VICT-) VICTORIAN DAIRY IND AUTHORITY.
(CSLC-) CSL LTD.

PI Reynolds EC, Obrien-Simpson NW, Slakeski N;

PP WPI; 1999-024043/02.

Immunogenic composition containing new Porphyromonas gingivalis peptides
- and related antibodies, useful for treatment, prevention and diagnosis
of periodontal disease.

PS Claim 1; Page 49; 70pp; English.

CC Peptides AAW83070-122 are derived from the 300 kD multiprotein complex of
cysteine proteinases and adhesins. The 300kDa multiprotein complex is the
major virulence factor of Porphyromonas gingivalis. The specification
describes a composition for raising an immune response against P.
gingivalis which contains at least one of the present peptides. The
products are used to treat and prevent P. gingivalis infection
(periodontal disease). Antibodies are used for passive immunisation, and

CC as diagnostic reagents for P. gingivalis in subgingival plaque. The
 XX peptides are used to detect P. gingivalis specific antibodies in serum
 CC

SQ Sequence 21 AA;

Query Match 100.0%; Score 118; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
 |||||
 DB 1 FNGGISLANYTGHGSETAWGT 21

RESULT 2

AAW83125

ID AAW83125 standard; protein; 493 AA.

XX

AC AAW83125;

XX 03-MAR-1999 (first entry)

XX

DE PrtIIR45 Arginine specific proteinase fragment.

XX

XX 300 kD multiprotein complex; cysteine proteinase; adhesin;
 KW virulence factor; immune response; P. gingivalis infection;
 KW periodontal disease; passive immunisation; subgingival plaque.

XX

OS Porphyromonas gingivalis.

XX

PN WO9849192-A1.

XX

PD 05-NOV-1998.

XX

PF 30-APR-1998; 98WO-AU000311.

XX

PR 30-APR-1997; 97AU-00006528.

XX

XX (UYME) UNIV MELBOURNE.

PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.

XX

XX (CSLC-) CSL LTD.

XX

PI Reynolds EC, Obrien-Simpson NM, Slakeski N;

XX

XX WPI; 1999-024043/02.

XX

DR Immunogenic composition containing new Porphyromonas gingivalis peptides

XX

PT - and related antibodies, useful for treatment, prevention and diagnosis

XX

PS of periodontal disease.

XX

PS Example 1; Fig 8; 70pp; English.

XX

XX The present sequence represents a fragment of the PrtIIR45 Arginine
 CC specific proteinase of Porphyromonas gingivalis. The specification
 CC describes peptides derived from the 300 kD multiprotein complex of
 CC cysteine proteinases and adhesins. The 300kDa multiprotein complex is the
 CC major virulence factor of P. gingivalis. The specification describes a
 CC composition for raising an immune response against P. gingivalis which
 CC contains at least one of the present peptides. The products are used to
 CC treat and prevent P. gingivalis infection (periodontal disease).

CC Antibodies are used for passive immunisation, and as diagnostic reagents
 CC for P. gingivalis in subgingival plaque. The peptides are used to detect
 CC P. gingivalis specific antibodies in serum

XX

SQ Sequence 493 AA;

Query Match 100.0%; Score 118; DB 2; Length 493;
 Best Local Similarity 100.0%; Pred. No. 2.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
 |||||
 DB 199 FNGGISLANYTGHGSETAWGT 219

RESULT 3

AAW83124

ID AAW83124 standard; protein; 507 AA.

XX

AC AAW83124;

XX

DT 03-MAR-1999 (first entry)

XX

XX PrtIIR50 Arginine specific proteinase fragment.

XX

XX 300 kD multiprotein complex; cysteine proteinase; adhesin;
 KW virulence factor; immune response; P. gingivalis infection;
 KW periodontal disease; passive immunisation; subgingival plaque.

XX

OS Porphyromonas gingivalis.

XX

PN WO9849192-A1.

XX

PD 05-NOV-1998.

XX

PF 30-APR-1998; 98WO-AU000311.

XX

PR 30-APR-1997; 97AU-00006528.

XX

XX (UYME) UNIV MELBOURNE.

PA

PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.

XX

PA (CSLC-) CSL LTD.

XX

PI Reynolds EC, Obrien-Simpson NM, Slakeski N;

XX

XX WPI; 1999-024043/02.

XX

Immunogenic composition containing new Porphyromonas gingivalis peptides

PT - and related antibodies, useful for treatment, prevention and diagnosis

PT of periodontal disease.

XX

XX Example 1; Fig 8; 70pp; English.

XX

XX The present sequence represents a fragment of the PrtIIR50 Arginine
 CC specific proteinase of Porphyromonas gingivalis. The specification
 CC describes peptides derived from the 300 kD multiprotein complex of
 CC cysteine proteinases and adhesins. The 300kDa multiprotein complex is the
 CC major virulence factor of P. gingivalis. The specification describes a
 CC composition for raising an immune response against P. gingivalis which
 CC contains at least one of the present peptides. The products are used to
 CC treat and prevent P. gingivalis infection (periodontal disease).

CC Antibodies are used for passive immunisation, and as diagnostic reagents
 CC for P. gingivalis in subgingival plaque. The peptides are used to detect
 CC P. gingivalis specific antibodies in serum

XX

SQ Sequence 507 AA;

Query Match 100.0%; Score 118; DB 2; Length 507;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
 |||||
 DB 199 FNGGISLANYTGHGSETAWGT 219

RESULT 4

AAW24786

ID AAW24786 standard; protein; 1706 AA.

XX

AC AAW24786;

XX

DT 17-OCT-2003 (revised)

XX

DT 25-NOV-1997 (first entry)

XX

DE PrtR antigenic protein complex.

XX	Periodontal disease; cell surface protein; thiol protease; endopeptidase; PrtR; PrtR45; PrtR44; PrtR15; PrtR17; PrtR27; haemagglutinin; adhesin; therapy; diagnosis; vaccine; antigen.	
XX	Porphyromonas gingivalis; strain W50.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..227
FT		/label= Pro-pro_peptide
FT	Cleavage-site	227..228
FT	Protein	228..719
FT		/label= PrtR45
FT		/note= "45 kDa Arg-specific thiol protease"
FT	Cleavage-site	719..720
FT	Protein	720..1138
FT		/label= PrtR44
FT		/note= "44 kDa adhesin"
FT	Cleavage-site	1138..1139
FT	Protein	1139..1273
FT		/label= PrtR15
FT		/note= "15 kDa adhesin"
FT	Cleavage-site	1273..1274
FT	Protein	1274..1431
FT		/label= PrtR17
FT		/note= "17 kDa adhesin"
FT	Cleavage-site	1431..1432
FT	Protein	1432..1706
FT		/label= PrtR27
FT		/note= "27 kDa adhesin"
XX		
FN	W09716542-Al.	
XX		
XX	09-MAY-1997.	
PD		
XX		
PF	30-OCT-1996; 96WO-AU000673.	
XX		
PR	30-OCT-1995; 95AU-00006275.	
XX		
PA	(UYME) UNIV MELBOURNE.	
PA	(VICT-) VICTORIAN DAIRY IND AUTHORITY.	
XX		
PI	Reynolds EC, Bhogal PS, Slakeski N;	
XX		
DR	WPI; 1997-272112/24.	
DR	N-PSDB; AAT78850.	
XX		
PT	New antigenic protein complex from Porphyromonas gingivalis - comprising Arg- and lys- specific thiol endo-peptidase(s), used in the detection, prevention and treatment of periodontal disease.	
XX		
PS	Example 1; Fig 8b; 68pp; English.	
XX		
CC	A PrtR-PrtRk cell surface protein of Porphyromonas ginivalis (PG) comprises a 300 kDa complex composed a 45 kDa arginine-specific thiol protease and 44, 15, 17 and 27 kDa adhesins encoded by the prtR gene (AAT78850), and a 148 kDa lysine-specific thiol protease and 39, 15 and 44 kDa adhesins (see AAT24787) encoded by the prtK gene (AAT78851). A claimed antigenic complex comprises at least one multimeric protein complex of PrtR and PrtK each containing at least one adhesin domain, the complex having a mol.wt. of over 200 kDa, and preferably comprises all 9 proteins of the PrtR-PrtK complex (see also AAT24780-85). It can be used in a claimed composition to elicit an immune response directed against PG, and in a claimed method of reducing the prospect of PG infection and/or severity of disease. Antibodies directed against the complex are claimed for use in treating PG infection. Unlike whole PG cells or other previously prepared antigens based on fimbriae or the capsule, the PrtR-PrtK complex or component parts are safe and effective antigens. (Updated on 17-OCT-2003 to standardise OS field)	
XX		
SQ	Sequence 1706 AA;	

Query Match 100.0%; Score 118; DB 2; Length 1706;

[illegible]

CC prepubertal periodontitis, generalized prepubertal periodontitis,
 CC localized juvenile periodontitis, generalized juvenile periodontitis,
 CC rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 CC abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.
 XX
 SQ Sequence 1706 AA;

Query Match 100.0%; Score 118; DB 9; Length 1706;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNGGISLANYTGHGSETAWGT 21
 DB 426 FNGGISLANYTGHGSETAWGT 446

RESULT 6
 ADZ10403
 ID ADZ10403 standard; protein; 1706 AA.

XX AC ADZ10403;

XX 16-JUN-2005 (first entry)

XX P. gingivalis arginine -specific protease ArgI SEQ ID 186.

XX Vaccine; enzyme; therapy; diagnosis; periodontal disease;
 KW antiinflammatory; mouth disease; gingivitis; inflammation; mouth disease;
 KW endocarditis; cardiant; cardiovascular disease; inflammation;
 KW bacterial urinary tract infection; antibacterial; uropathic;
 KW genitourinary disease; infection; osteomyelitis; osteopathic;
 KW musculoskeletal disease; antigen.

XX Porphyromonas gingivalis.

XX WO2005019249-A2.

XX 03-MAR-2005.

XX 10-AUG-2004; 2004WO-US025778.

XX 15-AUG-2003; 2003US-0495589P.

XX (UYFL) UNIV FLORIDA.

XX Progulske-Fox A, Hillman JD, Handfield M;

XX WPI; 2005-214225/22.

XX N-PSDB; ADZ10235.

XX New isolated immunogenic polypeptide, useful for diagnosing, treating,
 PT preventing, or ameliorating diseases and infections caused by
 PT Porphyromonas gingivalis e.g. localized prepubertal periodontitis.

XX Claim 1; SEQ ID NO 186; 73pp; English.

XX The invention relates to an isolated immunogenic Porphyromonas gingivalis
 CC polypeptide (I) comprising (a) a polypeptide at least about 5 contiguous
 CC amino acids of an amino acid sequence appearing as ADZ10384 - ADZ10443,
 CC ADZ10445 - ADZ10478 or ADZ10480-ADZ10571, (b) the polypeptide of (a) and
 CC a heterologous polypeptide or (c) an amino acid sequence selected from
 CC ADZ10384 - ADZ10443, ADZ10445 - ADZ10478 or ADZ10480-ADZ10571. Also
 CC included are an isolated polynucleotide (comprising a sequence that
 CC encodes (I) (a), (b) or (c)), an expression control sequence operably
 CC linked to the polynucleotide above (or the polynucleotide above and a
 CC heterologous polynucleotide), an expression vector comprising the
 CC polynucleotide, a host cell comprising the expression vector, an antibody
 CC (antibody fragment, or single-chain antibody) that specifically binds to
 CC the polypeptide, a composition comprising the antibody (antibody
 CC fragment, or single-chain antibody) and a pharmaceutical carrier,
 CC treating or preventing a disease/infection caused by Porphyromonas

CC gingivalis, a composition comprising the isolated immunogenic polypeptide
 CC (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 CC response in an animal, detecting the presence of a first P. gingivalis
 CC polynucleotide in a test sample, detecting the presence of a P.
 CC gingivalis antibody in a test sample, detecting the presence of P.
 CC gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting
 CC P. gingivalis infection in a subject and detecting P. gingivalis in a
 CC subject. The immunogenic polypeptide is useful for the diagnosis,
 CC treatment, prevention, and amelioration of diseases and infections caused
 CC by P. gingivalis. It is useful for treating or preventing localized
 CC prepubertal periodontitis, generalized prepubertal periodontitis,
 CC localized juvenile periodontitis, generalized juvenile periodontitis,
 CC rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 CC abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.

XX Sequence 1706 AA;

Query Match 100.0%; Score 118; DB 9; Length 1706;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNGGISLANYTGHGSETAWGT 21

DB 426 FNGGISLANYTGHGSETAWGT 446

RESULT 7
 ADZ10427

ID ADZ10427 standard; protein; 1706 AA.

XX AC ADZ10427;

XX 16-JUN-2005 (first entry)

XX P. gingivalis arginine specific protease ArgI, SEQ ID 210.

XX Vaccine; enzyme; therapy; diagnosis; periodontal disease;
 KW antiinflammatory; mouth disease; gingivitis; inflammation; mouth disease;
 KW endocarditis; cardiant; cardiovascular disease; inflammation;
 KW bacterial urinary tract infection; antibacterial; uropathic;
 KW genitourinary disease; infection; osteomyelitis; osteopathic;
 KW musculoskeletal disease; antigen.

XX Porphyromonas gingivalis.

XX WO2005019249-A2.

XX 03-MAR-2005.

XX 10-AUG-2004; 2004WO-US025778.

XX 15-AUG-2003; 2003US-0495589P.

XX (UYFL) UNIV FLORIDA.

XX Progulske-Fox A, Hillman JD, Handfield M;

XX WPI; 2005-214225/22.

XX N-PSDB; ADZ10256.

XX New isolated immunogenic polypeptide, useful for diagnosing, treating,
 PT preventing, or ameliorating diseases and infections caused by
 PT Porphyromonas gingivalis e.g. localized prepubertal periodontitis.

XX Claim 1; SEQ ID NO 210; 73pp; English.

XX The invention relates to an isolated immunogenic Porphyromonas gingivalis
 CC polypeptide (I) comprising (a) a polypeptide at least about 5 contiguous
 CC amino acids of an amino acid sequence appearing as ADZ10384 - ADZ10443,
 CC ADZ10445 - ADZ10478 or ADZ10480-ADZ10571, (b) the polypeptide of (a) and
 CC a heterologous polypeptide or (c) an amino acid sequence selected from

CC ADZ10384 - ADZ10443, ADZ10445 - ADZ10478 or ADZ10480-ADZ10571. Also
 CC included are an isolated polynucleotide (comprising a sequence that
 CC encodes (i) (a), (b) or (c)), an expression control sequence operably
 CC linked to the polynucleotide above (or the polynucleotide above and a
 CC heterologous polynucleotide), an expression vector comprising the
 CC polynucleotide, a host cell comprising the expression vector, an antibody
 CC (antibody fragment, or single-chain antibody) that specifically binds to
 CC the polypeptide, a composition comprising the antibody (antibody
 CC fragment, or single-chain antibody) and a pharmaceutical carrier,
 CC treating or preventing a disease/infection caused by Porphyromonas
 CC gingivalis, a composition comprising the isolated immunogenic polypeptide
 CC (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 CC response in an animal, detecting the presence of a first P. gingivalis
 CC polynucleotide in a test sample, detecting the presence of a P.
 CC gingivalis antibody in a test sample, detecting the presence of a P.
 CC gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting
 CC P. gingivalis infection in a subject and detecting P. gingivalis in a
 CC subject. The immunogenic polypeptide is useful for the diagnosis,
 CC treatment, prevention, and amelioration of diseases and infections caused
 CC by P. gingivalis. It is useful for treating or preventing localized
 CC prepubertal periodontitis, generalized prepubertal periodontitis,
 CC localized juvenile periodontitis, generalized juvenile periodontitis,
 CC rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 CC abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.
 XX
 SQ Sequence 1706 AA;

Query Match 100.0%; Score 118; DB 9; Length 1706;
 Best Local Similarity 100.0%; Pred. No. 1.le-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNGGISLANYTGHGSETAWGT 21
 |||||
 Db 426 FNGGISLANYTGHGSETAWGT 446

RESULT 8

ID ADZ10456
 ADZ10456 standard; protein; 1706 AA.

XX ADZ10456;

XX 16-JUN-2005 (first entry)

DE P. gingivalis arginine -specific protease ArgI SEQ ID 239.

XX Vaccine; enzyme; therapy; diagnosis; periodontal disease;
 KW antiinflammatory; mouth disease; gingivitis; inflammation; mouth disease;
 KW endocarditis; cardiant; cardiovascular disease; inflammation;
 KW bacterial urinary tract infection; antibacterial; uropathic;
 KW genitourinary disease; infection; osteomyelitis; osteopathic;
 KW musculoskeletal disease; antigen.

XX Porphyromonas gingivalis.

PN WO2005019249-A2.

XX 03-MAR-2005.

PF 10-AUG-2004; 2004WO-US025778.

XX 15-AUG-2003; 2003US-0495589P.

XX (UYFL) UNIV FLORIDA.

XX Progulske-Fox A, Hillman JD, Handfield M;

XX WPI: 2005-214225/22.

XX N-PSDB; ADZ10287.

PT New isolated immunogenic polypeptide, useful for diagnosing, treating,

PT preventing, or ameliorating diseases and infections caused by
 PT Porphyromonas gingivalis e.g. localized prepubertal periodontitis.
 XX Claim 1; SEQ ID NO 239; 73pp; English.

CC The invention relates to an isolated immunogenic Porphyromonas gingivalis
 CC polypeptide (I) comprising (a) a polypeptide at least about 5 contiguous
 CC amino acids of an amino acid sequence appearing as ADZ10384 - ADZ10443,
 CC ADZ10445 - ADZ10478 or ADZ10480-ADZ10571, (b) the polypeptide of (a) and
 CC a heterologous polypeptide or (c) an amino acid sequence selected from
 CC ADZ10384 - ADZ10443, ADZ10445 - ADZ10478 or ADZ10480-ADZ10571. Also
 CC included are an isolated polynucleotide (comprising a sequence that
 CC encodes (i) (a), (b) or (c)), an expression control sequence operably
 CC linked to the polynucleotide above (or the polynucleotide above and a
 CC heterologous polynucleotide), an expression vector comprising the
 CC polynucleotide, a host cell comprising the expression vector, an antibody
 CC (antibody fragment, or single-chain antibody) that specifically binds to
 CC the polypeptide, a composition comprising the antibody (antibody
 CC fragment, or single-chain antibody) and a pharmaceutical carrier,
 CC treating or preventing a disease/infection caused by Porphyromonas
 CC gingivalis, a composition comprising the isolated immunogenic polypeptide
 CC (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 CC response in an animal, detecting the presence of a first P. gingivalis
 CC polynucleotide in a test sample, detecting the presence of a P.
 CC gingivalis antibody in a test sample, detecting the presence of a P.
 CC gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting
 CC P. gingivalis infection in a subject and detecting P. gingivalis in a
 CC subject. The immunogenic polypeptide is useful for the diagnosis,
 CC treatment, prevention, and amelioration of diseases and infections caused
 CC by P. gingivalis. It is useful for treating or preventing localized
 CC prepubertal periodontitis, generalized prepubertal periodontitis,
 CC localized juvenile periodontitis, generalized juvenile periodontitis,
 CC rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 CC abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.
 XX

SQ Sequence 1706 AA;

Query Match 100.0%; Score 118; DB 9; Length 1706;
 Best Local Similarity 100.0%; Pred. No. 1.le-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNGGISLANYTGHGSETAWGT 21
 |||||
 Db 426 FNGGISLANYTGHGSETAWGT 446

RESULT 9

AEB91466

ID AEB91466 standard; protein; 1706 AA.

XX AEB91466;

XX 20-OCT-2005 (first entry)

DE Microbial pathogen adhesin protein sequence, SEQ ID NO:176.

XX algorithm; adhesin; pharmaceutical; vaccine; drug screening;
 KW bordetella pertussis infection; antibacterial; pneumonia;
 KW antiinflammatory; respiratory-gen.; gastric ulcer; antilucer;
 KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.

OS Porphyromonas gingivalis.

XX WO2005076010-A2.

XX 18-AUG-2005.

XX 07-FEB-2005; 2005WO-IN000037.

XX 06-FEB-2004; 2004IN-DE000173.

PR 20-JUL-2004; 2004US-0589227P.

XX The present invention relates to a method of producing activated protein
 CC C (APC) in a medium comprising protein C. The method involves contacting
 CC the medium with an arginine-specific cysteine proteinase, referred as
 CC Rgpa (also referred as 95-kDa gingipain R (HRGP)) and RgpB (also referred
 CC as 50 kDa-gingipain R (RGP-2)) of bacterial origin. The composition and
 CC methods are useful in controlling blood coagulation or in treating
 CC septicaemia or disseminated intravascular coagulation in patients. The
 CC invention is also used in gene therapy. The present sequence is
 CC Porphyromonas gingivalis 95 kDa RgpA protein 50 kDa catalytic domain
 XX
 SQ Sequence 492 AA;

Query Match 96.6%; Score 114; DB 7; Length 492;
 Best Local Similarity 95.2%; Pred. No. 1e-08;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
 ||||| ||||| ||||| ||||| |||||
 Db 199 FNGGISLVNYTGHGSETAWGT 219

RESULT 12
 ADV16838
 ID ADV16838 standard; protein; 492 AA.
 XX
 AC ADV16838;
 XX
 XX
 XX
 DT 24-FEB-2005 (first entry)
 XX
 DE Human cysteine protease related protein, SEQ ID 3.

XX selectable marker; periodontal disease; angina; antiangiinal;
 KW cardiovascular disease; diabetes; antidiabetic; cysteine protease;
 KW enzyme.
 KW
 XX Homo sapiens.

XX WO2004106541-A1.

XX 09-DEC-2004.

XX 31-MAY-2004; 2004WO-JP007867.

XX 30-MAY-2003; 2003JP-00154070.

XX (KYUS-) KYUSHU TLO CO LTD.

XX Yamamoto K;

XX WPI; 2005-021305/02.

XX Novel periodontal disease marker for detecting periodontal disease,
 PT comprising cysteine protease derived from Porphyromonas gingivalis,
 PT useful for detecting periodontal disease, in saliva.
 XX

PS Claim 5; SEQ ID NO 3; 44pp; Japanese.

XX The invention relates to a novel periodontal disease marker for detecting
 CC periodontal disease. The marker comprises a cysteine protease derived
 CC from a Porphyromonas gingivalis microbe. The invention further comprises:
 CC a periodontal disease ward appearance kit, comprising a sampling tool,
 CC periodontal disease marker, synthetic substrate, buffer, and a reducing
 CC agent. The periodontal disease marker is useful for detecting and
 CC diagnosing periodontal disease from saliva. The periodontal disease
 CC marker provides reliable and convenient detection of periodontal disease,
 CC and provides simultaneous detection of angina and diabetes. This sequence
 CC represents a human cysteine protease related protein of the invention.
 XX

SQ Sequence 492 AA;

Query Match 96.6%; Score 114; DB 9; Length 492;
 Best Local Similarity 95.2%; Pred. No. 1e-08;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
 ||||| ||||| ||||| ||||| |||||
 Db 199 FNGGISLVNYTGHGSETAWGT 219

RESULT 13
 ABW02695
 ID ABW02695 standard; protein; 507 AA.
 XX
 AC ABW02695;
 XX

DT 11-MAR-2004 (first entry)
 XX
 DE Porphyromonas gingivalis 50 kDa RgpB protein.
 XX

KW Activated protein C; APC; arginine-specific cysteine proteinase; Rgpa;
 KW 95-kDa gingipain R; HRGP; RgpB; 50 kDa-gingipain R; blood coagulation;
 KW RGP-2; septicaemia; disseminated intravascular coagulation; gene therapy;
 KW anticoagulant; thrombolytic; antibacterial; immunosuppressive; enzyme.
 XX
 OS Porphyromonas gingivalis.

XX US6627193-B1.

XX 30-SEP-2003.

XX 13-JAN-2000; 2000US-00482500.

XX 13-JAN-1999; 99US-0115869P.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

XX Travis J, Imamura T, Potempa J;

XX WPI; 2003-874308/81.

XX Producing activated protein C in blood, useful for controlling blood
 PT coagulation, comprises contacting the blood with an amount of an arginine
 PT -specific cysteine proteinase derived from Porphyromonas gingivalis.
 XX

PS Claim 1; SEQ ID NO 3; 0pp; English.

XX The present invention relates to a method of producing activated protein
 CC C (APC) in a medium comprising protein C. The method involves contacting
 CC the medium with an arginine-specific cysteine proteinase, referred as
 CC Rgpa (also referred as 95-kDa gingipain R (HRGP)) and RgpB (also referred
 CC as 50 kDa-gingipain R (RGP-2)) of bacterial origin. The composition and
 CC methods are useful in controlling blood coagulation or in treating
 CC septicaemia or disseminated intravascular coagulation in patients. The
 CC invention is also used in gene therapy. The present sequence is
 CC Porphyromonas gingivalis 50 kDa RgpB protein
 XX

SQ Sequence 507 AA;

Query Match 96.6%; Score 114; DB 7; Length 507;
 Best Local Similarity 95.2%; Pred. No. 1.1e-08;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
 ||||| ||||| ||||| ||||| |||||
 Db 199 FNGGISLVNYTGHGSETAWGT 219

RESULT 14
 AAR70186
 ID AAR70186 standard; protein; 737 AA.
 XX
 AC AAR70186;
 XX

DT 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)

```

XX DE Arg-gingipain-1.
XX KW Arg-gingipain-1; gingivalis; periodontal disease; vaccine;
XX KW arginine-specific protease.
XX OS Porphyromonas gingivalis.
XX PH Key Location/Qualifiers
XX FT Protein 228..737
XX FT /label= Mat_protein
XX PN WO9507286-A1.
XX XX
XX PD 16-MAR-1995.
XX PF 09-SEP-1994; 94WO-US010283.
XX PR 10-SEP-1993; 93US-00119361.
XX PR 21-OCT-1993; 93US-00141324.
XX PR 24-JUN-1994; 94US-00265441.
XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX PI Travis J, Potempa J, Barr PJ, Pavloff N;
XX PI WPI; 1995-123373/16.
XX DR N-PSDB; AAQ83484.
XX CC DNA encoding Arg-gingipain proteins - used to develop prods. for
XX CC detection, treatment and prevention of periodontal disease.
XX PS Disclosure; Page 63-66; 89pp; English.
XX CC A low mol.wt. arginine-specific gingipain (AG-1) and high mol.wt. AG (AG-
XX CC 2) were isolated from P. gingivalis strains H66 (ATCC 33277) and W50
XX CC (ATCC 53973). The sequences of the proteins were used to design PCR
XX CC primers and probes to isolate AG DNA. Lambda DASH and lambda ZAP
XX CC libraries were screened with a probe based on amino acids 11-22 of the AG
XX CC protein to obtain DNA encoding AG-1 (AAQ83484) and AG-2 (AAQ83489).
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 737 AA;
XX Query Match 96.6%; Score 114; DB 2; Length 737;
XX Best Local Similarity 95.2%; Pred. No. 1.7e-08;
XX Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 FNGGISLANYTGHGSETAWGT 21
XX Db ||||| ||||| ||||| ||||| |||||
XX 426 FNGGISLVNYTGHGSETAWGT 446
XX RESULT 15
XX AAQ34846
XX ID AAW34846 standard; protein; 737 AA.
XX AC AAW34846;
XX DT 03-JUN-1998 (first entry)
XX DE Arg-gingipain high molecular weight polypeptide sequence.
XX KW Arg-specific gingipain protease; gingivalis; periodontal disease;
XX KW vaccine; infection.
XX OS Porphyromonas gingivalis.
XX PH Key Location/Qualifiers
XX FT Protein 1..227
XX FT /note= "precursor protein"
XX FT Protein 228..737
XX FT /note= "mature Arg-gingipain"
XX FT

```

```

XX PN WO9734629-A1.
XX PD 25-SEP-1997.
XX PF 21-MAR-1997; 97WO-US004635.
XX PR 22-MAR-1996; 96US-0013945P.
XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX PA (MORE-) MOREHOUSE SCHOOL MEDICINE.
XX PI Potempa J, Travis J, Genco C;
XX DR WPI; 1997-479993/44.
XX DR N-PSDB; AAT93873.
XX PT Porphyromonas gingivalis Arg-specific gingipain protease peptide(s) -
XX PT useful for protecting animals and humans from gingivalis and periodontal
XX PS diseases.
XX PS Disclosure; Page 58-60; 95pp; English.
XX CC The present sequence represents an arginine-specific protease of
XX CC Porphyromonas gingivalis. The following peptides, derived from Arg- and
XX CC Lys-specific high molecular weight proteases, offer protection against
XX CC infection: YTVVYRDGK IKEGLTATTE DDGATGNHE YCWEKYTAGS VSPKVC (I);
XX CC YTPVEEKQNG RMIVIVAKKY (II); QLPFIFDVAC VNGDFLFSMP CFAEALMRAQ (III);
XX CC GEPNPQOPVS NLTAATGGQK VTLKWDASTK (IV); GNHEYCVVEVK YTAGVSPKVC KDVTV (V);
XX CC RMFNYPFGR YTPVEEKQNG (VI); TFAGFEDTYK RMFNYPFGR (VII); DYTVVYRDG
XX CC TKIKEGLTAT TFEEDGATG NMEYCVCKY TAGVSPKVC (VIII); YTVVYRDGT KIKEGLTATTF
XX CC EEDG (IX); RDGTKIKEGL TATTEEDGV ATGN (X); KIKEGLTATT FEEDGATGN HEY (XI)
XX CC ; KWDAPNGTNP PNP ENPNP PGTTLSE (XII); and YTPVEEKENG RMIVIVAKKY
XX CC (XIII). They are used in vaccines to protect animals, including humans,
XX CC from gingivitis and/or periodontal diseases
XX SQ Sequence 737 AA;
XX Query Match 96.6%; Score 114; DB 2; Length 737;
XX Best Local Similarity 95.2%; Pred. No. 1.7e-08;
XX Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 FNGGISLANYTGHGSETAWGT 21
XX Db ||||| ||||| ||||| ||||| |||||
XX 426 FNGGISLVNYTGHGSETAWGT 446
XX Search completed: August 25, 2006, 17:57:47
XX Job time : 216.372 secs

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:04:46 ; Search time 34.6744 Seconds
(without alignments)
53.012 Million cell updates/sec

Title: US-10-387-977-1

Perfect score: 118

Sequence: 1 FNGGISLANTYGHGSETAWGT 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC Celerra_SID33/ptodata/2/iaa/5 COMB.pep.*
- 2: /EMC Celerra_SID33/ptodata/2/iaa/6 COMB.pep.*
- 3: /EMC Celerra_SID33/ptodata/2/iaa/7 COMB.pep.*
- 4: /EMC Celerra_SID33/ptodata/2/iaa/H COMB.pep.*
- 5: /EMC Celerra_SID33/ptodata/2/iaa/2CTUS COMB.pep.*
- 6: /EMC Celerra_SID33/ptodata/2/iaa/RE COMB.pep.*
- 7: /EMC Celerra_SID33/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	118	100.0	1706	2	US-09-066-330-10
2	114	96.6	492	2	US-09-482-500A-2
3	114	96.6	507	2	US-09-482-500A-3
4	114	96.6	737	1	US-08-119-361-5
5	114	96.6	737	2	US-08-336-308A-4
6	114	96.6	737	2	US-08-822-324-4
7	114	96.6	737	2	US-09-490-931-4
8	114	96.6	1477	2	US-09-482-500A-1
9	114	96.6	1687	1	US-08-570-311-29
10	114	96.6	1704	2	US-08-336-308A-10
11	114	96.6	1704	2	US-08-822-324-6
12	114	96.6	1704	2	US-09-490-931-10
13	82	69.5	1732	1	US-08-570-311-10
14	82	69.5	1732	1	US-08-353-485-10
15	82	69.5	1732	2	US-09-066-330-11
16	75	63.6	942	1	US-08-141-324-14
17	75	63.6	942	1	US-08-541-902-14
18	75	63.6	1358	1	US-08-570-311-27
19	72	61.0	509	2	US-08-822-324-8
20	57	48.3	382	2	US-09-134-000C-3738
21	54	45.8	207	2	US-09-149-476-516
22	54	45.8	232	2	US-09-149-476-633
23	54	45.8	252	2	US-09-149-476-332
24	49.5	41.9	47	2	US-10-375-913-7
25	49.5	41.9	47	2	US-11-110-001-7
26	49.5	41.9	47	2	US-11-110-002-7

27	48	40.7	244	2	US-09-107-532A-5529	Sequence 5529, Ap
28	48	40.7	556	2	US-09-134-000C-5850	Sequence 5850, Ap
29	48	40.7	614	1	US-08-295-814E-12	Sequence 12, Appl
30	48	40.7	614	1	US-08-543-881-2	Sequence 2, Appl
31	48	40.7	614	1	US-08-291-299-2	Sequence 2, Appl
32	48	40.7	614	1	US-08-291-299-8	Sequence 8, Appl
33	48	40.7	614	2	US-09-343-361-12	Sequence 12, Appl
34	48	40.7	614	2	US-09-919-039-378	Sequence 378, Appl
35	48	40.7	614	5	PCT-US94-00119-2	Sequence 2, Appl
36	48	40.7	614	5	PCT-US95-10579-2	Sequence 2, Appl
37	48	40.7	614	5	PCT-US95-10579-8	Sequence 8, Appl
38	48	40.7	634	2	US-09-949-016-7681	Sequence 7681, Ap
39	48	40.7	1012	1	US-08-475-891A-4	Sequence 4, Appl
40	48	40.7	1025	1	US-08-567-375-4	Sequence 4, Appl
41	48	40.7	1025	1	US-08-587-680A-4	Sequence 4, Appl
42	48	40.7	1026	2	US-09-623-551-18	Sequence 18, Appl
43	47	39.8	932	2	US-09-248-796A-19128	Sequence 19128, A
44	45	38.1	243	2	US-09-328-352-7058	Sequence 7058, Ap
45	45	38.1	404	2	US-09-543-681A-6702	Sequence 6702, Ap

ALIGNMENTS

RESULT 1

US-09-066-330-10
; Sequence 10, Application US/09066330A
; Patent No. 6511666
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: Bhogal, Peter S.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: Reynolds
; CURRENT APPLICATION NUMBER: US/09/066.330A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: PN 6275
; EARLIER FILING DATE: 1995-10-30
; EARLIER APPLICATION NUMBER: PCT/AU96/00673
; EARLIER FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 10
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-066-330-10

Query Match 100.0%; Score 118; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNGGISLANTYGHGSETAWGT 21
Db 426 FNGGISLANTYGHGSETAWGT 446

RESULT 2

US-09-482-500A-2
; Sequence 2, Application US/09482500A
; Patent No. 6627193
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Inamura, Takahisa
; APPLICANT: Potempa, Jan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION
; FILE REFERENCE: 235.00160101
; CURRENT APPLICATION NUMBER: US/09/482.500A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/115,869
; PRIOR FILING DATE: 1999-01-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-482-500A-2

Query Match          96.6%; Score 114; DB 2; Length 492;
Best Local Similarity 95.2%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
    ||||| ||||| ||||| |||||
Db 199 FNGGISLVNYTGHGSETAWGT 219

RESULT 3
US-09-482-500A-3
; Sequence 3, Application US/09482500A
; Patent No. 6627193
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Imamura, Takahisa
; APPLICANT: Potempa, Jan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION
; FILE REFERENCE: 235.00160101
; CURRENT APPLICATION NUMBER: US/09/482,500A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/115,869
; PRIOR FILING DATE: 1999-01-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-482-500A-3

Query Match          96.6%; Score 114; DB 2; Length 507;
Best Local Similarity 95.2%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
    ||||| ||||| ||||| |||||
Db 199 FNGGISLVNYTGHGSETAWGT 219

RESULT 4
US-08-119-361-5
; Sequence 5, Application US/08119361
; Patent No. 5523390
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis Arginine-specific Proteinase
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,361
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 21-93C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-308A-4
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33878
; REFERENCE/DOCKET NUMBER: 21-93
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-499-8080
; TELEFAX: 303-499-8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-119-361-5

Query Match          96.6%; Score 114; DB 1; Length 737;
Best Local Similarity 95.2%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
    ||||| ||||| ||||| |||||
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 5
US-08-336-308A-4
; Sequence 4, Application US/08336308A
; Patent No. 6017532
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis
; TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,308A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/119,361
; FILING DATE: 10-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,441
; FILING DATE: 24-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 21-93C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-308A-4
```

```
Query Match          96.6%; Score 114; DB 2; Length 737;
Best Local Similarity 95.2%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLVNYTGHGSETAWGT 21
||||| ||||||| |||||||
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 6
US-08-822-324-4
; Sequence 4, Application US/08822324
; Patent No. 6129917
; GENERAL INFORMATION:
; APPLICANT: Potempa, Jan S.
; APPLICANT: Travis, James
; APPLICANT: Genco, Caroline A.
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,945
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 103-95 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 498-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-822-324-4

Query Match          96.6%; Score 114; DB 2; Length 737;
Best Local Similarity 95.2%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLVNYTGHGSETAWGT 21
||||| ||||||| |||||||
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 7
US-09-490-931-4
; Sequence 4, Application US/09490931
; Patent No. 6274718
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
```

```
; TITLE OF INVENTION: Porphyromonas gingivalis
; TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,931
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/336,308
; FILING DATE:
; APPLICATION NUMBER: US 08/265,441
; FILING DATE: 24-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 21-93C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-490-931-4

Query Match          96.6%; Score 114; DB 2; Length 737;
Best Local Similarity 95.2%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLVNYTGHGSETAWGT 21
||||| ||||||| |||||||
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 8
US-09-482-500A-1
; Sequence 1, Application US/09482500A
; Patent No. 6627193
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Imamura, Takahisa
; APPLICANT: Potempa, Jan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION
; FILE REFERENCE: 235.00160101
; CURRENT APPLICATION NUMBER: US/09/482,500A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/115,869
; PRIOR FILING DATE: 1999-01-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1477
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-482-500A-1

Query Match          96.6%; Score 114; DB 2; Length 1477;
Best Local Similarity 95.2%; Pred. No. 5.7e-08;
```

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLVNYTGHGSETAWGT 21
Db 199 FNGGISLVNYTGHGSETAWGT 219

RESULT 9

US-08-570-311-29
; Sequence 29, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, GUYLAINE
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C3
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1687 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-570-311-29

Query Match 96.6%; Score 114; DB 1; Length 1687;
Best Local Similarity 95.2%; Pred. No. 6.7e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLVNYTGHGSETAWGT 21
Db 407 FNGGISLVNYTGHGSETAWGT 427

RESULT 10

US-08-336-308A-10
; Sequence 10, Application US/08336308A
; Patent No. 6017532
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis
; ARGinine-specific Proteinase Coding Sequences
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,308A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/119,361
; FILING DATE: 10-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,441
; FILING DATE: 24-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 21-93C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-308A-10

Query Match 96.6%; Score 114; DB 2; Length 1704;
Best Local Similarity 95.2%; Pred. No. 6.8e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLVNYTGHGSETAWGT 21
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 11

US-08-822-324-6
; Sequence 6, Application US/08822324
; Patent No. 6129917
; GENERAL INFORMATION:
; APPLICANT: Potempa, Jan S.
; APPLICANT: Travis, James
; APPLICANT: Genco, Caroline A.
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder

```

; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,945
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 103-95 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 488-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-822-324-6
;
Query Match 96.6%; Score 114; DB 2; Length 1704;
Best Local Similarity 95.2%; Pred. No. 6.8e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 12
US-09-490-931-10
; Sequence 10, Application US/09490931
; Patent No. 6274718
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis
; TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,931
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/336,308
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,441
; FILING DATE: 24-JUN-1994

; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,945
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 103-95 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 488-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-822-324-6
;
Query Match 96.6%; Score 114; DB 2; Length 1704;
Best Local Similarity 95.2%; Pred. No. 6.8e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 13
US-08-570-311-10
; Sequence 10, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 10:

```

```

; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,945
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 103-95 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 488-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-822-324-6
;
Query Match 96.6%; Score 114; DB 2; Length 1704;
Best Local Similarity 95.2%; Pred. No. 6.8e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 12
US-09-490-931-10
; Sequence 10, Application US/09490931
; Patent No. 6274718
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis
; TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,931
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/336,308
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,441
; FILING DATE: 24-JUN-1994

; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,945
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 103-95 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 488-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-822-324-6
;
Query Match 96.6%; Score 114; DB 2; Length 1704;
Best Local Similarity 95.2%; Pred. No. 6.8e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLVNYTGHGSETAWGT 21
Db 426 FNGGISLVNYTGHGSETAWGT 446

RESULT 13
US-08-570-311-10
; Sequence 10, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 10:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10

Query Match 69.5%; Score 82; DB 1; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.0032;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NGGISLANYTGHGSETAW 19
Db 433 NTGVSFANYTAHGETAW 450

RESULT 14
US-08-353-485-10
Sequence 10, Application US/08353485
Patent No. 5830710

GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C2
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-10

Query Match 69.5%; Score 82; DB 1; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.0032;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NGGISLANYTGHGSETAW 19

Db 433 NTGVSFANYTAHGETAW 450
RESULT 15
US-09-066-330-11
Sequence 11, Application US/09066330A
Patent No. 6511666
GENERAL INFORMATION:
APPLICANT: Reynolds, Eric C.
APPLICANT: Bhogal, Peter S.
APPLICANT: Slakeski, Nada
TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
FILE REFERENCE: Reynolds
CURRENT APPLICATION NUMBER: US/09/066.330A
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: PN 6275
EARLIER FILING DATE: 1995-10-30
EARLIER APPLICATION NUMBER: PCT/AU96/00673
EARLIER FILING DATE: 1996-10-30
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 11
LENGTH: 1732
TYPE: PRT
ORGANISM: Porphyromonas gingivalis
US-09-066-330-11

Query Match 69.5%; Score 82; DB 2; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.0032;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NGGISLANYTGHGSETAW 19
Db 433 NTGVSFANYTAHGETAW 450

Search completed: August 25, 2006, 18:06:29
Job time : 35.6744 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:21:12 ; Search time 121.116 Seconds
(without alignments)
80.315 Million cell updates/sec

Title: US-10-387-977-1
Perfect score: 118
Sequence: 1 FNGGISLANYTGHGSETAWGT 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA Main:*
1: /EMC_Celerra_SID83/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /EMC_Celerra_SID83/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
3: /EMC_Celerra_SID83/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
4: /EMC_Celerra_SID83/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
5: /EMC_Celerra_SID83/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
6: /EMC_Celerra_SID83/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	21	4	US-10-387-977-1
2	118	100.0	491	4	US-10-387-977-100
3	118	100.0	507	4	US-10-387-977-99
4	118	100.0	736	4	US-10-387-977-105
5	118	100.0	1706	4	US-10-229-066-10
6	118	100.0	1706	5	US-10-915-002-186
7	118	100.0	1706	5	US-10-915-002-210
8	118	100.0	1706	5	US-10-915-002-221
9	118	100.0	1706	5	US-10-915-002-239
10	118	100.0	1706	6	US-11-052-554A-176
11	82	69.5	22	4	US-10-387-977-2
12	82	69.5	509	4	US-10-387-977-101
13	82	69.5	1731	5	US-10-915-002-176
14	82	69.5	1731	5	US-10-915-002-192
15	82	69.5	1731	5	US-10-915-002-211
16	82	69.5	1731	5	US-10-915-002-222
17	82	69.5	1732	4	US-10-229-066-11
18	58.5	49.6	450	4	US-10-369-493-8448
19	58.5	49.6	475	6	US-11-087-099-5183
20	58.5	49.6	475	6	US-11-188-298-15792
21	54	45.8	207	3	US-09-809-391-516
22	54	45.8	207	3	US-09-882-171-516
23	54	45.8	207	4	US-10-164-861-516
24	54	45.8	207	6	US-11-144-947-516
25	54	45.8	232	3	US-09-809-391-633
26	54	45.8	232	3	US-09-882-171-633
27	54	45.8	232	4	US-10-164-861-633

28	54	45.8	232	6	US-11-144-947-633	Sequence 633, App
29	54	45.8	252	3	US-09-809-391-332	Sequence 332, App
30	54	45.8	252	3	US-09-882-171-332	Sequence 332, App
31	54	45.8	252	4	US-10-164-861-332	Sequence 332, App
32	54	45.8	252	6	US-11-144-947-332	Sequence 332, App
33	54	45.8	1123	5	US-10-915-002-245	Sequence 245, App
34	50	42.4	491	6	US-11-188-298-4899	Sequence 4899, App
35	50	42.4	817	4	US-10-437-963-114827	Sequence 114827, App
36	49.5	41.9	47	4	US-10-375-913-7	Sequence 7, Appli
37	49.5	41.9	47	6	US-11-110-001-7	Sequence 7, Appli
38	49.5	41.9	47	6	US-11-110-002-7	Sequence 7, Appli
39	49.5	41.9	47	6	US-11-235-009-7	Sequence 7, Appli
40	49.5	41.9	323	3	US-09-925-301-952	Sequence 952, App
41	49	41.5	132	4	US-10-425-115-324154	Sequence 324154, App
42	48	40.7	263	6	US-11-096-568A-11199	Sequence 11199, A
43	48	40.7	351	6	US-11-096-568A-11198	Sequence 11198, A
44	48	40.7	416	4	US-10-369-493-20445	Sequence 20445, A
45	48	40.7	525	4	US-10-424-599-260318	Sequence 260318, App

ALIGNMENTS

RESULT 1

US-10-387-977-1
; Sequence 1, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR FILING DATE: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-1
Query Match 100.0%; Score 118; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FNGGISLANYTGHGSETAWGT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 1 FNGGISLANYTGHGSETAWGT 21
RESULT 2
US-10-387-977-100
; Sequence 100, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977

```
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-100

Query Match      100.0%; Score 118; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||
Db 199 FNGGISLANYTGHGSETAWGT 219

RESULT 3
US-10-387-977-99
; Sequence 99, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-99

Query Match      100.0%; Score 118; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 3.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||
Db 199 FNGGISLANYTGHGSETAWGT 219

RESULT 4
US-10-387-977-105
; Sequence 105, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
```

```
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-105

Query Match      100.0%; Score 118; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||
Db 428 FNGGISLANYTGHGSETAWGT 448

RESULT 5
US-10-229-066-10
; Sequence 10, Application US/10229066
; Publication No. US20030157637A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: Bhogal, Peter S.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: Reynolds
; CURRENT APPLICATION NUMBER: US/10/229,066
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/066,330
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: PN 6275
; PRIOR FILING DATE: 1995-10-30
; PRIOR APPLICATION NUMBER: PCT/AU96/00673
; PRIOR FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-229-066-10

Query Match      100.0%; Score 118; DB 4; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||
Db 426 FNGGISLANYTGHGSETAWGT 446

RESULT 6
US-10-915-002-186
; Sequence 186, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Progulskie-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS,
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
```



```

; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 186
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-185

```

```
Query Match      100.0%; Score 118; DB 5; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy		1	FNGGISLANYTGHGSETAWGT	21
Dd		426	FNGGISLANYTGHGSETAWGT	446

```

RESULT 7
US-10-915-002-210
; Sequence 210, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, AND
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 210
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-210

```

```
Query Match      100.0%; Score 118; DB 5; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 FNGGISLANYTGHGSETAWGT 21
|||
Dp 426 FNGGISLANYTGHGSETAWGT 446

RESULT 8
 US-10-915-002-221
 ; Sequence 221, Application US/10915002
 ; Publication No. US20060078950A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Progulskie-Fox, Ann
 ; APPLICANT: Hillman, Jeffrey D.
 ; APPLICANT: Handfield, Martin
 ; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDES
 ; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
 ; TITLE OF INVENTION: PERIODONTAL DISEASES
 ; FILE REFERENCE: 02-042
 ; CURRENT APPLICATION NUMBER: US/10/915,002
 ; CURRENT FILING DATE: 2004-08-10
 ; NUMBER OF SEQ ID NOS: 354
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 221
 ; LENGTH: 1706
 ; TYPE: PRT
 ; ORGANISM: Porphyromonas gingivalis
 US-10-915-002-221

```
Query Match      100.0%; Score 118; DB 5; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
```

```

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNGGISLANYTGHGSETAWCT 21
    |||||
Db 426 FNGGISLANYTGHGSETAWCT 446

```

```

RESULT 9
US-10-915-002-239
; Sequence 239, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Progleuke-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDES
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT,
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 239
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-239

```

Query Match	100.0%;	Score 118;	DB 5;	Length 1706;
Best Local Similarity	100.0%;	Pred. NO. 1.4e-08;		
Matches 21;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps 0;

Qy	1	21
Qy	1	21
Db	426	446

```

RESULT 10
US-11-052-554A-176
; Sequence 176, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 176
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis W83
US-11-052-554A-176

```

Query Match	100.0%	Score 118;	DB 6;	Length 1706;
Best Local Similarity	100.0%;	pred. NO. 1.4e-08;		
Matches 21:	Conservative	0;	Mismatches	0;
	Indels	0;		Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
|||
Dp 426 FNGGISLANYTGHGSETAWGT 446

RESULT 11
US-10-387-977-2
: Sequence 2. Application US/10387977

```
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-2

Query Match          69.5%; Score 82; DB 4; Length 22;
Best Local Similarity 77.8%; Pred. No. 3.7e-05;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 NGGISLANYTGHGSETAW 19
Db      2 NTGVSFANYTAHGSETAW 19

RESULT 12
US-10-387-977-101
; Sequence 101, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-101

Query Match          69.5%; Score 82; DB 4; Length 509;
Best Local Similarity 77.8%; Pred. No. 0.0011;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 NGGISLANYTGHGSETAW 19
Db      205 NTGVSFANYTAHGSETAW 222

RESULT 13
US-10-915-002-176
; Sequence 176, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 176
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-176

Query Match          69.5%; Score 82; DB 5; Length 1731;
Best Local Similarity 77.8%; Pred. No. 0.0041;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 NGGISLANYTGHGSETAW 19
Db      433 NTGVSFANYTAHGSETAW 450

RESULT 14
US-10-915-002-192
; Sequence 192, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 192
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-192

Query Match          69.5%; Score 82; DB 5; Length 1731;
Best Local Similarity 77.8%; Pred. No. 0.0041;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 NGGISLANYTGHGSETAW 19
Db      433 NTGVSFANYTAHGSETAW 450

RESULT 15
US-10-915-002-211
; Sequence 211, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
```

```
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 211
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-211
```

```
Query Match      69.5%; Score 82; DB 5; Length 1731;
Best Local Similarity 77.8%; Pred. No. 0.0041;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      2  NGGISLANYTGHGSETAW 19
      | | | | | | | | | | | | | | |
Db      433 NTGVSFANYTAGSETAW 450
```

```
Search completed: August 25, 2006, 18:26:06
Job time : 122.116 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:22:07 ; Search time 20.0233 Seconds
(without alignments)
71.760 Million cell updates/sec

Title: US-10-387-977-1

Perfect score: 118

Sequence: 1 FNGGISLANYTHGHSETAWG 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 247503 seqs, 68422524 residues

Total number of hits satisfying chosen parameters: 247503

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New.*

- 1: /EMC_Celerra_SID33/ptodata/2/pubpaa/US09_NEW_PUB.pdb.*
- 2: /EMC_Celerra_SID33/ptodata/2/pubpaa/US06_NEW_PUB.pdb.*
- 3: /EMC_Celerra_SID33/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
- 4: /EMC_Celerra_SID33/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
- 5: /EMC_Celerra_SID33/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
- 6: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10_NEW_PUB.pdb.*
- 7: /EMC_Celerra_SID33/ptodata/2/pubpaa/US11_NEW_PUB.pdb.*
- 8: /EMC_Celerra_SID33/ptodata/2/pubpaa/US60_NEW_PUB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	48	40.7	263	7 US-11-056-355B-69319	Sequence 69319, A
2	48	40.7	351	7 US-11-056-355B-69318	Sequence 69318, A
3	48	40.7	614	7 US-11-300-928-62	Sequence 62, Appl
4	47	39.8	389	6 US-10-449-902-33007	Sequence 33007, A
5	46	39.0	948	6 US-10-449-902-44448	Sequence 44448, A
6	45.5	38.6	266	6 US-10-449-902-37409	Sequence 37409, A
7	45	38.1	506	6 US-10-953-349-10213	Sequence 10213, A
8	45	38.1	548	6 US-10-953-349-10212	Sequence 10212, A
9	45	38.1	563	6 US-10-449-902-46522	Sequence 46522, A
10	45	38.1	574	6 US-10-953-349-10211	Sequence 10211, A
11	44	37.3	151	6 US-10-449-902-38955	Sequence 38955, A
12	44	37.3	393	6 US-10-449-902-41675	Sequence 41675, A
13	44	37.3	751	7 US-11-330-403-8183	Sequence 8183, A
14	43	36.4	249	6 US-10-374-780A-1540	Sequence 1540, Ap
15	42	35.6	192	6 US-10-449-902-41740	Sequence 41740, A
16	42	35.6	301	7 US-11-056-355B-66160	Sequence 66160, A
17	42	35.6	312	6 US-10-953-349-278	Sequence 278, App
18	42	35.6	320	7 US-11-174-307B-3740	Sequence 3740, Ap
19	42	35.6	320	7 US-11-056-355B-66159	Sequence 66159, A
20	42	35.6	355	6 US-10-953-349-277	Sequence 277, App
21	42	35.6	408	6 US-10-953-349-276	Sequence 276, App
22	42	35.6	636	6 US-10-449-902-44935	Sequence 44935, A
23	42	35.6	661	7 US-11-056-355B-47185	Sequence 47185, A
24	42	35.6	680	7 US-11-056-355B-47184	Sequence 47184, A
25	42	35.6	730	6 US-10-449-902-51873	Sequence 51873, A

ALIGNMENTS

RESULT 1

US-11-056-355B-69319
; Sequence 69319, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:

; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69319
; LENGTH: 263
; TYPE: prt
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(263)
; OTHER INFORMATION: Ceres Seq. ID no. 15222289
US-11-056-355B-69319

Query Match 40.7%; Score 48; DB 7; Length 263;
Best Local Similarity 39.1%; Pred No. 8.3;
Matches 9; Conservative 5; Mismatches 5; Indels 4; Gaps 1;

Qy 2 NGGISLAN-----YTHGHSETAWG 20
|||:| |||:|:|
Db 189 NGGVPLTRQGSLSYTGVTQIGVG 211

RESULT 2

US-11-056-355B-69318
; Sequence 69318, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:

; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR FILING DATE: 2004-02-13
US-11-056-355B-69319

Sequence 47183, A
Sequence 45698, A
Sequence 1654, Ap
Sequence 30389, A
Sequence 47941, A
Sequence 7181, Ap
Sequence 7180, Ap
Sequence 3948, Ap
Sequence 37993, A
Sequence 82889, A
Sequence 82888, A
Sequence 82887, A
Sequence 122, App
Sequence 14014, A
Sequence 5058, Ap
Sequence 55224, A
Sequence 55228, A
Sequence 38064, A
Sequence 58732, A
Sequence 58731, A

```

/ NUMBER OF SEQ ID NOS: 119966
/
/ SEQ ID NO 69318
/ LENGTH: 351
/ TYPE: prt
/ ORGANISM: Triticum aestivum
/
/ FEATURE:
/
/ NAME/KEY: peptide
/ LOCATION: (1)..(351)
/
/ OTHER INFORMATION: Ceres Seq
US-11-056-355B-69318

```

Query Match	40.7%	Score 48;	DB 7;	Length 351;
Best Local Similarity	39.1%	Pred. No. 11;		
Matches	9;	Conservative	5;	Mismatches 5;
				Indels 4;
				Gaps 1;

QY 2 NGGISLAN---YTHGSETAWG 20
|||:| |||:|:|
Db 277 NGGVPLTROGSLYTGTOIGYG 299

RESULT 3

```

US-11-300-928-62
; Sequence 62, Application US/11300928
; Publication NO. US20060166277A1
; GENERAL INFORMATION:
; APPLICANT: Karumanchi, S. Ananth
; APPLICANT: Sukhatme, Vikas P.
; TITLE OF INVENTION: Nucleic Acids and Polypeptides Useful For Diagnosing and Treating
; TITLE OF INVENTION: Complications of Pregnancy
; FILE REFERENCE: 01948/108002
; CURRENT APPLICATION NUMBER: US/11/300,928
; CURRENT FILING DATE: 2005-12-15
; PRIOR APPLICATION NUMBER: US 60/636,275
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 62
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-300-928-62

```

Query Match	40.7%	Score 48	DB 7	Length 614
Best Local Similarity	53.3%	Pred. No. 20		
Matches 8	Conservative	2	Mismatches 5	Indels 0
Matches	Conservative	2	Mismatches 5	Indels 0
Matches	Conservative	2	Mismatches 5	Indels 0

```
Qy      5 ISLANYTGHGSETAW 19
          ::||| ||| |||
Db      94 VALGOYTSOGSVTAW 108
```

RESULT 4

RES001 4
 US-10-449-902-33007
 ; Sequence 33007, Application US/10449902
 ; Publication No. US20060123505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: National Institute of Agrobiological Sciences.
 ; APPLICANT: Bio-oriented Technology Research Advancement Institution.
 ; APPLICANT: The Institute of Physical and Chemical Research.
 ; APPLICANT: Foundation for Advancement of International Science.
 ; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
 ; FILE REFERENCE: MOA-A0205Y1-US
 ; CURRENT APPLICATION NUMBER: US/10/449,902
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: JP 2002-203269
 ; PRIOR FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: JP 2002-383870
 ; PRIOR FILING DATE: 2002-12-11
 ; NUMBER OF SEQ ID NOS: 56791
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 33007
 ; LENGTH: 389
 ; TYPE: PRT

Query Match	38.6%;	Score 45.5;	DB 6;	Length 266;
Best Local Similarity	50.0%;	Pred. No. 20;		

; ORGANISM: Oryza sativa
US-10-449-902-33007

Query Match 39.8%; Score 47; DB 6; Length 389;
Best Local Similarity 42.9%; Pred. No. 18;
Matches 9; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

Qy 2 NGGISLANY--TGHGSETAWG 20
: || || || || : ||
db 17 SGGYGOONYRPOGHGAASQWG 37

RESULT 5

US-10-449-902-44448
; Sequence 44448, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020531-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 44448
; LENGTH: 948
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-44448

Query Match 39.0%; Score 46; DB 6; Length 948;
Best Local Similarity 58.3%; Pred. No. 64;
Matches 7: Conservative 1; Mismatches 4; Indels 0; Gaps 0

```

Qy      10 YTHGSETAWGT 21
         | | | | : | | |
Dh      133 YAGHGNEEDWAT 144

```

RESULT 6

```

US-10-449-902-37409
; Sequence 37409, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37409
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-37409

```

Query Match	38.6%;	Score 45.5;	DB 6;	Length 266;
Best Local Similarity	50.0%;	Pred. No. 20;		

```
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 3 GGISLANYTGHGSETAW 20
   ||:|:|:|:|:|
Db 179 GGSAPASY-GYGSNSGWS 195

RESULT 7
US-10-953-349-10213
; Sequence 10213, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10213
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10213

Query Match 38.1%; Score 45; DB 6; Length 506;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 0;

Qy 4 GISLANYTGHGSETAW 19
   ||:|:|:|:|
Db 123 GLSMENRANGSBEAW 138

RESULT 8
US-10-953-349-10212
; Sequence 10212, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10212
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10212

Query Match 38.1%; Score 45; DB 6; Length 548;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 0;

Qy 4 GISLANYTGHGSETAW 19
   ||:|:|:|:|
Db 165 GLSMENRANGSBEAW 180

RESULT 9
US-10-449-902-46522
; Sequence 46522, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
```

```
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46522
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-46522

Query Match 38.1%; Score 45; DB 6; Length 563;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 10; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

Qy 2 NG--GISLANYTGHGSETAW 19
   ||:|:|:|:|
Db 542 NGLAGYSLDGTSHGPQAAW 561

RESULT 10
US-10-953-349-10211
; Sequence 10211, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10211
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10211

Query Match 38.1%; Score 45; DB 6; Length 574;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GISLANYTGHGSETAW 19
   ||:|:|:|:|
Db 191 GLSMENRANGSBEAW 206

RESULT 11
US-10-449-902-38955
; Sequence 38955, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38955
; LENGTH: 151
```

```
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-38955

Query Match      37.3%; Score 44; DB 6; Length 151;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHS 15
Db 108 FNGGFLAIVAGHAA 122

RESULT 12
US-10-449-902-41675
; Sequence 41675, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41675
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-41675

Query Match      37.3%; Score 44; DB 6; Length 393;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHS2TA 18
Db 68 FSGSGGSRGCGFSES 85

RESULT 13
US-11-330-403-8183
; Sequence 8183, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 8183
; LENGTH: 751
; TYPE: PRT
; ORGANISM: ASPERGILLUS NIDULANS FGSC A4
US-11-330-403-8183

Query Match      37.3%; Score 44; DB 7; Length 751;
Best Local Similarity 42.9%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHS2TA 21
Db 153 FNLAARLARYTGNQYADWAT 173
```

```
RESULT 14
US-10-374-780A-1540
; Sequence 1540, Application US/10374780A
; Publication No. US20060162006A9
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1540
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Orthologous to GI322
US-10-374-780A-1540

Query Match      36.4%; Score 43; DB 6; Length 249;
Best Local Similarity 47.1%; Pred. No. 45;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 ISLANYTGHS2TA 21
Db 31 LSLVNYIAANGEGAWNT 47

RESULT 15
US-10-449-902-41740
; Sequence 41740, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
```

```
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41740
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-41740

Query Match      35.6%; Score 42; DB 6; Length 192;
Best Local Similarity 40.0%; Pred. No. 49;
Matches 8; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY      2  NGGISLANYTGHGSETAWGT 21
      || | | | | | | | | | |
Db      50  NGNKGNGNGNGNGLNYWGT 69

Search completed: August 25, 2006, 18:26:52
Job time : 21.0233 secs
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:05:37 ; Search time 417.558 Seconds
(without alignments)
76.794 Million cell updates/sec

Title: US-10-387-977-1
Perfect score: 118
Sequence: 1 FNGGISLANYTGHSETAWGT 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/paa/US066 COMB.pap.*
2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066 COMB.pap.*
3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073 COMB.pap.*
4: /EMC_Celerra_SIDS3/ptodata/2/paa/US074 COMB.pap.*
5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075 COMB.pap.*
6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076 COMB.pap.*
7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077 COMB.pap.*
8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078 COMB.pap.*
9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079 COMB.pap.*
10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080 COMB.pap.*
11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081 COMB.pap.*
12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082 COMB.pap.*
13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083 COMB.pap.*
14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084 COMB.pap.*
15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085 COMB.pap.*
16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086 COMB.pap.*
17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087 COMB.pap.*
18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088 COMB.pap.*
19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089 COMB.pap.*
20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090 COMB.pap.*
21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091 COMB.pap.*
22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092 COMB.pap.*
23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093 COMB.pap.*
24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094 COMB.pap.*
25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095 COMB.pap.*
26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096 COMB.pap.*
27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097 COMB.pap.*
28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098 COMB.pap.*
29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099 COMB.pap.*
30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100 COMB.pap.*
31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101 COMB.pap.*
32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102 COMB.pap.*
33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103 COMB.pap.*
34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104 COMB.pap.*
35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105 COMB.pap.*
36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106 COMB.pap.*
37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107 COMB.pap.*
38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108 COMB.pap.*
39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109 COMB.pap.*
40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110 COMB.pap.*
41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111 COMB.pap.*
42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112 COMB.pap.*
43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113 COMB.pap.*

44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114 COMB.pap.*
45: /EMC_Celerra_SIDS3/ptodata/2/paa/US600 COMB.pap.*
46: /EMC_Celerra_SIDS3/ptodata/2/paa/US601 COMB.pap.*
47: /EMC_Celerra_SIDS3/ptodata/2/paa/US602 COMB.pap.*
48: /EMC_Celerra_SIDS3/ptodata/2/paa/US603 COMB.pap.*
49: /EMC_Celerra_SIDS3/ptodata/2/paa/US604 COMB.pap.*
50: /EMC_Celerra_SIDS3/ptodata/2/paa/US605 COMB.pap.*
51: /EMC_Celerra_SIDS3/ptodata/2/paa/US606 COMB.pap.*
52: /EMC_Celerra_SIDS3/ptodata/2/paa/US607 COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	118	100.0	21	24	US-09-423-056B-1	Sequence 1, Appli
2	118	100.0	21	24	US-09-423-056B-1	Sequence 1, Appli
3	118	100.0	21	33	US-10-387-977-1	Sequence 1, Appli
4	118	100.0	491	24	US-09-423-056B-100	Sequence 100, App
5	118	100.0	491	24	US-09-423-056B-100	Sequence 100, App
6	118	100.0	491	33	US-10-387-977-100	Sequence 100, App
7	118	100.0	507	24	US-09-423-056B-99	Sequence 99, Appl
8	118	100.0	507	24	US-09-423-056B-99	Sequence 99, Appl
9	118	100.0	507	33	US-10-387-977-99	Sequence 105, App
10	118	100.0	736	24	US-09-423-056B-105	Sequence 21158, A
11	118	100.0	736	27	US-09-791-537-21158	Sequence 105, App
12	118	100.0	736	33	US-10-387-977-105	Sequence 49353, A
13	118	100.0	1526	27	US-09-791-537-49353	Sequence 100403,
14	118	100.0	1706	27	US-09-791-537-113656	Sequence 113656,
15	118	100.0	1706	27	US-09-791-537-113656	Sequence 10, Appl
16	118	100.0	1706	32	US-10-229-066-10	Sequence 186, App
17	118	100.0	1706	33	US-10-915-002-186	Sequence 210, App
18	118	100.0	1706	39	US-10-915-002-210	Sequence 221, App
19	118	100.0	1706	39	US-10-915-002-221	Sequence 233, App
20	118	100.0	1706	39	US-10-915-002-239	Sequence 176, App
21	118	100.0	1706	40	US-11-052-554A-176	Sequence 186, App
22	118	100.0	1706	49	US-60-495-589-186	Sequence 210, App
23	118	100.0	1706	49	US-60-495-589-210	Sequence 221, App
24	118	100.0	1706	49	US-60-495-589-221	Sequence 239, App
25	118	100.0	1706	49	US-60-495-589-239	Sequence 29, Appl
26	114	96.6	394	25	US-09-543-696B-29	Sequence 49386, A
27	114	96.6	435	27	US-09-791-537-49386	Sequence 1, Appli
28	114	96.6	435	29	US-09-979-624-1	Sequence 27, Appl
29	114	96.6	492	25	US-09-543-696B-27	Sequence 28, Appl
30	114	96.6	492	25	US-09-543-696B-27	Sequence 28, Appl
31	114	96.6	508	25	US-09-543-696B-28	Sequence 26, Appl
32	114	96.6	508	25	US-09-543-696B-28	Sequence 26, Appl
33	114	96.6	736	25	US-09-543-696B-26	Sequence 55763, A
34	114	96.6	736	27	US-09-791-537-55763	Sequence 94999, A
35	114	96.6	736	27	US-09-791-537-94999	Sequence 4, Appli
36	114	96.6	737	1	PCT-US94-10283-4	Sequence 5, Appli
37	114	96.6	737	12	US-08-265-441-5	Sequence 4, Appli
38	114	96.6	737	12	US-08-265-441-5	Sequence 4, Appli
39	114	96.6	737	25	US-09-543-696B-4	Sequence 4, Appli
40	114	96.6	737	25	US-09-543-696B-4	Sequence 4, Appli
41	114	96.6	737	29	US-09-927-018-4	Sequence 65556, A
42	114	96.6	991	27	US-09-791-537-65556	Sequence 18, Appl
43	114	96.6	1185	12	US-08-265-441-18	Sequence 29, Appl
44	114	96.6	1687	21	US-09-174-517-29	
45	114	96.6				

ALIGNMENTS

RESULT 1
US-09-423-056-1
; Sequence 1, Application US/09423056
; GENERAL INFORMATION:

```
; APPLICANT: Reynolds, Eric C.
; APPLICANT: O'Brien-Simpson, Neil M.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: Synthetic peptide constructs for the diagnosis and
; TITLE OF INVENTION: treatment of periodontitis associated with
; TITLE OF INVENTION: Porphyromonas gingivalis
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/423,056
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-423-056-1

Query Match      100.0%; Score 118; DB 24; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNGGISLANYTGHGSETAWGT 21
Db 1 FNGGISLANYTGHGSETAWGT 21
|||||
|||||

RESULT 2
US-09-423-056B-1
; Sequence 1, Application US/09423056B
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000300
; CURRENT APPLICATION NUMBER: US/09/423,056B
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-423-056B-1

Query Match      100.0%; Score 118; DB 24; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNGGISLANYTGHGSETAWGT 21
Db 1 FNGGISLANYTGHGSETAWGT 21
|||||
|||||

RESULT 3
US-10-387-977-1
; Sequence 1, Application US/10387977
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
```

```
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-1

Query Match      100.0%; Score 118; DB 33; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNGGISLANYTGHGSETAWGT 21
Db 1 FNGGISLANYTGHGSETAWGT 21
|||||
|||||

RESULT 4
US-09-423-056-100
; Sequence 100, Application US/09423056
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: O'Brien-Simpson, Neil M.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: Synthetic peptide constructs for the diagnosis and
; TITLE OF INVENTION: treatment of periodontitis associated with
; TITLE OF INVENTION: Porphyromonas gingivalis
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/423,056
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 100
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-423-056-100

Query Match      100.0%; Score 118; DB 24; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNGGISLANYTGHGSETAWGT 21
Db 199 FNGGISLANYTGHGSETAWGT 219
|||||
|||||

RESULT 5
US-09-423-056B-100
; Sequence 100, Application US/09423056B
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000300
; CURRENT APPLICATION NUMBER: US/09/423,056B
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
```

```
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-423-056B-100

Query Match      100.0%; Score 118; DB 24; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||||
Db 199 FNGGISLANYTGHGSETAWGT 219
   |||||||

RESULT 6
US-10-387-977-100
; Sequence 100, Application US/10387977
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; FILE REFERENCE: 52928200301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-100

Query Match      100.0%; Score 118; DB 33; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||||
Db 199 FNGGISLANYTGHGSETAWGT 219
   |||||||

RESULT 7
US-09-423-056B-99
; Sequence 99, Application US/09423056
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: O'Brien-Simpson, Neil M.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: Synthetic peptide constructs for the diagnosis and
; TITLE OF INVENTION: treatment of periodontitis associated with
; TITLE OF INVENTION: Porphyromonas gingivalis
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/423,056
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
US-10-387-977-99

Query Match      100.0%; Score 118; DB 24; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||||
Db 199 FNGGISLANYTGHGSETAWGT 219
   |||||||

RESULT 8
US-09-423-056B-99
; Sequence 99, Application US/09423056B
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; FILE REFERENCE: 52928200300
; CURRENT APPLICATION NUMBER: US/09/423,056B
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-423-056B-99

Query Match      100.0%; Score 118; DB 24; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
   |||||||
Db 199 FNGGISLANYTGHGSETAWGT 219
   |||||||

RESULT 9
US-10-387-977-99
; Sequence 99, Application US/10387977
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-99
```

us-10-387-977-1.rapm

Mon Aug 28 11:30:41 2006

```

; Sequence 105, Application US/10387977
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200301
; CURRENT APPLICATION NUMBER: US/10387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-10-387-977-105

Query Match 100.0%; Score 118; DB 33; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
Db 199 FNGGISLANYTGHGSETAWGT 219

RESULT 10
US-09-423-056B-105
; Sequence 105, Application US/09423056B
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200300
; CURRENT APPLICATION NUMBER: US/09/423,056B
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-09-423-056B-105

Query Match 100.0%; Score 118; DB 24; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
Db 428 FNGGISLANYTGHGSETAWGT 448

RESULT 11
US-09-791-537-21158
; Sequence 21158, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21158
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-09-791-537-21158

Query Match 100.0%; Score 118; DB 27; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
Db 428 FNGGISLANYTGHGSETAWGT 448

RESULT 12
US-10-387-977-105
; Sequence 105, Application US/10387977
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200301
; CURRENT APPLICATION NUMBER: US/10387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-10-387-977-105

Query Match 100.0%; Score 118; DB 33; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
Db 199 FNGGISLANYTGHGSETAWGT 219

RESULT 13
US-09-791-537-49353
; Sequence 49353, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 49353
; LENGTH: 1526
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-09-791-537-49353

Query Match 100.0%; Score 118; DB 27; Length 1526;
Best Local Similarity 100.0%; Pred. No. 4.8e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
Db 426 FNGGISLANYTGHGSETAWGT 446

RESULT 14
US-09-791-537-100403
; Sequence 100403, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210

```

; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 100403
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-791-537-100403

Query Match 100.0%; Score 118; DB 27; Length 1706;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 426 FNGGISLANYTGHGSETAWGT 446

RESULT 15

US-09-791-537-113656
; Sequence 113656, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113656
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-791-537-113656

Query Match 100.0%; Score 118; DB 27; Length 1706;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 426 FNGGISLANYTGHGSETAWGT 446

Search completed: August 25, 2006, 18:20:51
Job time : 418.558 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:06:46 ; Search time 26.3721 Seconds
(without alignments)
68.789 Million cell updates/sec

Title: US-10-387-977-1

Perfect score: 118

Sequence: 1 FNGGISLANYTGHSETAWGT 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 363143 seqs, 86385820 residues

Total number of hits satisfying chosen parameters: 363143

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA New:*

- 1: /EMC_Celerra_SID83/ptodata/2/paa/PCT_NEW_COMB.pap.*
- 2: /EMC_Celerra_SID83/ptodata/2/paa/US06_NEW_COMB.pap.*
- 3: /EMC_Celerra_SID83/ptodata/2/paa/US07_NEW_COMB.pap.*
- 4: /EMC_Celerra_SID83/ptodata/2/paa/US08_NEW_COMB.pap.*
- 5: /EMC_Celerra_SID83/ptodata/2/paa/US09_NEW_COMB.pap.*
- 6: /EMC_Celerra_SID83/ptodata/2/paa/US10_NEW_COMB.pap.*
- 7: /EMC_Celerra_SID83/ptodata/2/paa/US11_NEW_COMB.pap.*
- 8: /EMC_Celerra_SID83/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	118	100.0	1706	6	US-10-229-066A-21
2	82	69.5	1732	6	US-10-229-066A-24
3	60.5	51.3	433	1	PCT-US06-18535-26437
4	60.5	51.3	433	7	US-11-431-855-26437
5	53.5	45.3	468	1	PCT-US06-18535-26266
6	53.5	45.3	468	7	US-11-431-855-26266
7	50.5	42.8	471	1	PCT-US06-18535-26339
8	50.5	42.8	471	7	US-11-431-855-26339
9	50	42.4	491	1	PCT-US06-18535-19305
10	50	42.4	491	7	US-11-431-855-19305
11	48	40.7	614	7	US-11-431-708-3588
12	48	40.7	614	7	US-11-431-708-3590
13	48	40.7	614	7	US-11-431-708-3601
14	48	40.7	614	7	US-11-431-708-3609
15	48	40.7	614	7	US-11-431-708-3614
16	48	40.7	614	7	US-11-475-062-6827
17	48	40.7	614	7	US-11-475-062-6829
18	48	40.7	614	7	US-11-475-062-6840
19	48	40.7	614	7	US-11-475-062-6848
20	48	40.7	614	7	US-11-475-062-6853
21	44	37.3	1256	8	US-60-836-986-9889
22	43	36.4	90	7	US-11-366-965-1049
23	43	36.4	269	7	US-11-371-354-58591
24	43	36.4	269	8	US-60-836-986-22583
25	43	36.4	289	1	PCT-US06-18535-24456

26	43	36.4	289	7	US-11-431-855-24456	Sequence 24456, A
27	43	36.4	766	8	US-60-836-986-28711	Sequence 28711, A
28	42.5	36.0	1709	8	US-60-836-986-6642	Sequence 6642, Ap
29	42.5	36.0	1962	8	US-60-836-986-6643	Sequence 6643, Ap
30	42	35.6	322	7	US-11-371-354-72899	Sequence 72899, A
31	42	35.6	322	7	US-11-371-354-78411	Sequence 78411, A
32	42	35.6	322	8	US-60-836-986-6589	Sequence 6589, Ap
33	42	35.6	325	7	US-11-371-354-71467	Sequence 71467, A
34	42	35.6	325	8	US-60-836-986-6588	Sequence 6588, Ap
35	42	35.6	339	1	PCT-US06-18535-12487	Sequence 12487, A
36	42	35.6	339	7	US-11-431-855-12487	Sequence 12487, A
37	42	35.6	394	1	PCT-US06-18535-16446	Sequence 16446, A
38	42	35.6	394	7	US-11-431-855-16446	Sequence 16446, A
39	41.5	35.2	326	1	PCT-US06-18535-32019	Sequence 32019, A
40	41.5	35.2	326	7	US-11-431-855-32019	Sequence 32019, A
41	41.5	35.2	353	1	PCT-US06-18535-28955	Sequence 28955, A
42	41.5	35.2	353	7	US-11-431-855-28955	Sequence 28955, A
43	41.5	35.2	406	1	PCT-US06-18535-29315	Sequence 29315, A
44	41.5	35.2	406	7	US-11-431-855-29315	Sequence 29315, A
45	41.5	35.2	660	1	PCT-US06-16539-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-229-066A-21

; Sequence 21, Application US/10229066A

; GENERAL INFORMATION:

; APPLICANT: REYNOLDS, ERIC CHARLES

; APPLICANT: BHOGAL, PETER SINGH

; APPLICANT: SLAKESKI, NADA

; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE

; FILE REFERENCE: 4137-20

; CURRENT APPLICATION NUMBER: US/10/229,066A

; CURRENT FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: 09/066,330

; PRIOR FILING DATE: 1998-09-15

; PRIOR APPLICATION NUMBER: PCT/AU96/00673

; PRIOR FILING DATE: 1996-10-30

; PRIOR APPLICATION NUMBER: AU PN 6275

; PRIOR FILING DATE: 1995-10-30

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 3.3

; SEQ ID NO 21

; LENGTH: 1706

; TYPE: PRT

; ORGANISM: Porphyromonas gingivalis

US-10-229-066A-21

Query Match 100.0%; Score 118; DB 6; Length 1706;

Best Local Similarity 100.0%; Pred.No. 4.8e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qv 1 FNGGISLANYTGHSETAWGT 21

Db 426 FNGGISLANYTGHSETAWGT 446

RESULT 2

US-10-229-066A-24

; Sequence 24, Application US/10229066A

; GENERAL INFORMATION:

; APPLICANT: REYNOLDS, ERIC CHARLES

; APPLICANT: BHOGAL, PETER SINGH

; APPLICANT: SLAKESKI, NADA

; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE

; FILE REFERENCE: 4137-20

; CURRENT APPLICATION NUMBER: US/10/229,066A

; CURRENT FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: 09/066,330

; PRIOR FILING DATE: 1998-09-15

; PRIOR APPLICATION NUMBER: PCT/AU96/00673

; PRIOR FILING DATE: 1996-10-30
; PRIOR APPLICATION NUMBER: AU PN 6275
; PRIOR FILING DATE: 1995-10-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 24
; LENGTH: 1732
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-229-066A-24

Query Match 69.5%; Score 82; DB 6; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.00025;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NGGISLANYTGHGSETAW 19
Db 433 NTGVSFANYTAHGETAW 450
| : | | | | | | | | | |

RESULT 3
PCT-US06-18535-26437
; Sequence 26437, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26437
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R1808
PCT-US06-18535-26437

Query Match 51.3%; Score 60.5; DB 1; Length 433;
Best Local Similarity 57.1%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 4 GISLANYTGHGS---ETAWGT 21
Db 164 GIGFANYTGHGFMKPGPWGT 184
| : | | | | | | | | | |

RESULT 4
US-11-431-855-26437
; Sequence 26437, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26437
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R1808
US-11-431-855-26437

Query Match 51.3%; Score 60.5; DB 7; Length 433;
Best Local Similarity 57.1%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 4 GISLANYTGHGS---ETAWGT 21
Db 164 GIGFANYTGHGFMKPGPWGT 184
| : | | | | | | | | | |

RESULT 5

PCT-US06-18535-26266
; Sequence 26266, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)A
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26266
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
PCT-US06-18535-26266

Query Match 45.3%; Score 53.5; DB 1; Length 468;
Best Local Similarity 47.6%; Pred. No. 2.2;
Matches 10; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 4 GISLANYTGHGS---ETAWGT 21
Db 199 GVGFPANYTAHGGFMKPGPWGT 219
| : | | | | | | | | | |

RESULT 6
US-11-431-855-26266
; Sequence 26266, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26266
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
US-11-431-855-26266

Query Match 45.3%; Score 53.5; DB 7; Length 468;
Best Local Similarity 47.6%; Pred. No. 2.2;
Matches 10; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 4 GISLANYTGHGS---ETAWGT 21
Db 199 GVGFPANYTAHGGFMKPGPWGT 219
| : | | | | | | | | | |

RESULT 7
PCT-US06-18535-26339
; Sequence 26339, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)A
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26339
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
PCT-US06-18535-26339

Query Match 42.8%; Score 50.5; DB 1; Length 471;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 10; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 4 GISLANYTGHGS---ETAWG 20

```
Db      189 GIGFANYTAHGGFPKGVWG 208
      || ||||| || : ||
      || ||||| || : ||

RESULT 8
US-11-431-855-26339
; Sequence 26339, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26339
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
US-11-431-855-26339

Query Match      42.8%; Score 50.5; DB 7; Length 471;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 10; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Qy      4 GISLANYTGHGS---ETAWG 20
      || ||||| || : ||
      || ||||| || : ||

Db      189 GIGFANYTAHGGFPKGVWG 208

RESULT 9
PCT-US06-18535-19305
; Sequence 19305, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)A
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19305
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Bacteroides thetaiotaomicron VPI-5482
PCT-US06-18535-19305

Query Match      42.4%; Score 50; DB 1; Length 491;
Best Local Similarity 69.2%; Pred. No. 8.2;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      9 NYTGHGSETAWGT 21
      || ||||| || : ||
      || ||||| || : ||

Db      476 NWTGHGGETKSGT 488

RESULT 10
US-11-431-855-19305
; Sequence 19305, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19305
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Bacteroides thetaiotaomicron VPI-5482
US-11-431-855-19305

Query Match      42.4%; Score 50; DB 7; Length 491;
Best Local Similarity 69.2%; Pred. No. 8.2;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      9 NYTGHGSETAWGT 21
      || ||||| || : ||
      || ||||| || : ||

Db      476 NWTGHGGETKSGT 488

RESULT 11
US-11-431-708-3588
; Sequence 3588, Application US/11431708
; GENERAL INFORMATION:
; APPLICANT: RUBEN, Steven
; TITLE OF INVENTION: BREAST DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001599-ORD
; CURRENT APPLICATION NUMBER: US/11/431,708
; CURRENT FILING DATE: 2006-05-11
; NUMBER OF SEQ ID NOS: 8544
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3588
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-431-708-3588

Query Match      40.7%; Score 48; DB 7; Length 614;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      5 ISLANYTGHGSETAW 19
      : || || || || ||
      : || || || || ||

Db      94 VALGQYTSQGSVTAW 108

RESULT 12
US-11-431-708-3590
; Sequence 3590, Application US/11431708
; GENERAL INFORMATION:
; APPLICANT: RUBEN, Steven
; TITLE OF INVENTION: BREAST DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001599-ORD
; CURRENT APPLICATION NUMBER: US/11/431,708
; CURRENT FILING DATE: 2006-05-11
; NUMBER OF SEQ ID NOS: 8544
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3590
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-431-708-3590

Query Match      40.7%; Score 48; DB 7; Length 614;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      5 ISLANYTGHGSETAW 19
      : || || || || ||
      : || || || || ||

Db      94 VALGQYTSQGSVTAW 108

RESULT 13
US-11-431-708-3601
; Sequence 3601, Application US/11431708
; GENERAL INFORMATION:
; APPLICANT: RUBEN, Steven
; TITLE OF INVENTION: BREAST DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001599-ORD
; CURRENT APPLICATION NUMBER: US/11/431,708
; CURRENT FILING DATE: 2006-05-11
; NUMBER OF SEQ ID NOS: 8544
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3601
```



```

; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-431-708-3601

Query Match      40.7%; Score 48; DB 7; Length 614;
Best Local Similarity 53.3%; Pred. No. 21;
Matches      8; Conservative      2; Mismatches      5; Indels      0; Gaps      0;

QY      5 ISLANYTGHGSETAW 19
      ::||| ||| |||
Db      94 VALGQYTSQGSVTAW 108

RESULT 14
US-11-431-708-3609
; Sequence 3609, Application US/11431708
; GENERAL INFORMATION:
; APPLICANT: RUBEN, Steven
; TITLE OF INVENTION: BREAST DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CLO01599-ORD
; CURRENT APPLICATION NUMBER: US/11/431,708
; CURRENT FILING DATE: 2006-05-11
; NUMBER OF SEQ ID NOS: 8544
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3609
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-431-708-3609

Query Match      40.7%; Score 48; DB 7; Length 614;
Best Local Similarity 53.3%; Pred. No. 21;
Matches      8; Conservative      2; Mismatches      5; Indels      0; Gaps      0;

QY      5 ISLANYTGHGSETAW 19
      ::||| ||| |||
Db      94 VALGQYTSQGSVTAW 108

RESULT 15
US-11-431-708-3614
; Sequence 3614, Application US/11431708
; GENERAL INFORMATION:
; APPLICANT: RUBEN, Steven
; TITLE OF INVENTION: BREAST DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CLO01599-ORD
; CURRENT APPLICATION NUMBER: US/11/431,708
; CURRENT FILING DATE: 2006-05-11
; NUMBER OF SEQ ID NOS: 8544
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3614
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-431-708-3614

Query Match      40.7%; Score 48; DB 7; Length 614;
Best Local Similarity 53.3%; Pred. No. 21;
Matches      8; Conservative      2; Mismatches      5; Indels      0; Gaps      0;

QY      5 ISLANYTGHGSETAW 19
      ::||| ||| |||
Db      94 VALGQYTSQGSVTAW 108

Search completed: August 25, 2006, 18:21:51
Job time : 27.3721 secs

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 17:58:43 ; Search time 23.4419 Seconds
(without alignments)
86.194 Million cell updates/sec

Title: US-10-387-977-1
Perfect score: 118
Sequence: 1 FNGGISLVNTGHGSETAWGT 21
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	1526	2 S49763	gingipain R (EC 3.1.1.1)
2	114	96.6	991	2 I40229	arginyl endopeptidase
3	114	96.6	1704	2 A55426	gingipain R (EC 3.1.1.1)
4	82	69.5	1732	2 T30836	lysine-specific cy
5	53	44.9	234	2 S41406	arginine-specific
6	49.5	41.9	542	2 I39540	chitinase (EC 3.2.1.1)
7	48	40.7	612	2 T10727	protein kinase Xa2
8	48	40.7	614	2 S68236	betaine/GABA trans
9	48	40.7	614	2 A41757	betaine transport
10	48	40.7	677	2 A27286	levanase (EC 3.2.1.1)
11	48	40.7	813	2 T04313	protein kinase Xa2
12	48	40.7	996	2 T10725	protein kinase Xa2
13	48	40.7	1025	1 A57676	protein kinase Xa2
14	47	39.8	61	2 G83651	hypothetical prote
15	47	39.8	294	2 AE2695	conserved hypothet
16	46	39.8	294	2 E97477	hypothetical prote
17	46	39.0	404	2 E82012	sodium/glutamate s
18	46	39.0	404	2 F81240	sodium/glutamate s
19	46	39.0	1002	2 T09438	toXR-activated lip
20	46	39.0	1013	2 B82276	ToXR-activated gen
21	46	39.0	2136	2 B84651	hypothetical prote
22	45.5	38.6	760	2 B64817	probable membrane
23	45.5	38.6	760	2 B85589	hypothetical prote
24	45.5	38.6	760	2 C90739	hypothetical prote
25	45	38.1	130	2 S65682	elkp protein - Sta
26	45	38.1	212	2 T22437	hypothetical prote
27	45	38.1	248	2 C75140	hypothetical prote
28	45	38.1	307	2 AB2029	ribokinase [fmpor
29	45	38.1	548	2 T49948	hypothetical prote

30	45	38.1	1736	2 T00391	hypothetical prote
31	45	38.1	4639	1 A54794	dynein heavy chain
32	44.5	37.7	287	2 S65765	chitinase (EC 3.2.1.1)
33	44.5	37.7	328	2 AE0301	conserved hypothet
34	44	37.3	289	2 S23402	sperm surface prot
35	44	37.3	458	2 E70145	aminopeptidase I (
36	44	37.3	465	2 B69768	probable acid-CoA
37	44	37.3	614	2 A43390	gamma-aminobutyric
38	44	37.3	793	2 S59067	penton long fiber
39	44	37.3	2871	2 A55567	fibrillin I - bovi
40	43.5	36.9	281	2 S18245	xyLF protein - Pse
41	43.5	36.9	555	2 AC0210	pectate disacchari
42	43	36.4	169	2 F72532	hypothetical prote
43	43	36.4	205	2 A37329	antigen S - paper
44	43	36.4	254	2 D35957	probable 2-deoxy-D
45	43	36.4	301	2 B84533	hypothetical prote

ALIGNMENTS

RESULT 1

S49763

gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)

C;Species: Porphyromonas gingivalis

C;Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C;Accession: S49763

R;Aduse-Opoku, J.; Muir, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A.

submitted to the EMBL Data Library, November 1994

A;Description: Cloning, sequence analysis and expression in Escherichia coli of prpR1

A;Reference number: S49763

A;Accession: S49763

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1526 <ADU>

A;Cross-references: UNIPROT:Q51838; UNIPARC:UPI0000179912; EMBL:X82680

C;Genetics:

A;Gene: prpR1

C;Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 118; DB 2; Length 1526;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLVNTGHGSETAWGT 21
|||||
DB 426 FNGGISLVNTGHGSETAWGT 446

RESULT 2

I40229

arginyl endopeptidase - Porphyromonas gingivalis

C;Species: Porphyromonas gingivalis

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004

C;Accession: I40229

R;Okamoto, K.; Misumi, Y.; Kadowaki, T.; Yoneda, M.; Yamamoto, K.; Ikehara, Y.

Arch. Biochem. Biophys. 316, 917-925, 1995

A;Title: Structural characterization of argingipain, a novel arginine-specific cysteine

A;Reference number: I40229; MUID:95168884; PMID:7864651

A;Accession: I40229

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-991 <RES>

A;Cross-references: UNIPROT:P28784; UNIPARC:UPI000012829F; GB:D26470; NID:9927644; PIDN:

Query Match 96.6%; Score 114; DB 2; Length 991;
Best Local Similarity 95.2%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLVNTGHGSETAWGT 21
|||||
DB 426 FNGGISLVNTGHGSETAWGT 446

```
RESULT 3
A58426
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
N;Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP; R
C;Species: Porphyromonas gingivalis
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C;Accession: A55426; D53113
R;Pavloff, N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr, F
J. Biol. Chem. 270, 1007-1010, 1995
A;Title: Molecular cloning and structural characterization of the Arg-gingipain proteinase
A;Reference number: A55426; MUID:95138080; PMID:7836351
A;Accession: A55426
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1704 <PAV>
A;Cross-references: UNIPROT:Q51816; UNIPARC:UPI00000B7BC1; GB:U15282; NID:g557067; PIDN:
R;Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
J. Biol. Chem. 269, 406-411, 1994
A;Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolati
A;Reference number: A53113; MUID:94103245; PMID:8276827
A;Accession: D53113
A;Status: preliminary
A;Molecule type: protein
A;Residues: 228-249 <PIK>
A;Cross-references: UNIPARC:UPI00000B9226
A;Experimental source: H66
A;Note: sequence extracted from NCBI backbone (NCBIP:141694)
C;Keywords: cysteine proteinase; hydrolase

Query Match 96.6%; Score 114; DB 2; Length 1704;
Best Local Similarity 95.2%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNGGISLANYTGHGSETAWGT 21
||||| ||||| ||||| ||||| |||||
Db 426 FNGGISLWNYTGHGSETAWGT 446

RESULT 4
T30836
lysine-specific cysteine proteinase porphypain (EC 3.4.22.-) - Porphyromonas gingivalis
N;Alternate names: lysine-specific cysteine proteinase 1, 60K
C;Species: Porphyromonas gingivalis
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30837; T30526; A53113
R;Barkocy-Gallagher, G.A.; Han, N.; Patti, J.M.; Whitlock, J.; Progulski-Fox, A.; Lantz,
J. Bacteriol. 178, 2734-2741, 1996
A;Title: Analysis of the prtP gene encoding porphypain, a cysteine proteinase of Porphy
A;Reference number: Z20895; MUID:96213011; PMID:8631659
A;Accession: T30836
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1732 <BAR>
A;Cross-references: UNIPROT:Q51817; UNIPARC:UPI00000B93C0; EMBL:U42210; NID:gl314325; PI
R;Slakeski, N.; Cleal, S.M.; Reynolds, E.C.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20896
A;Accession: T30837
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-795,'I',797-1389,'N',1391-1478,'Y',1480-1732 <SLA>
A;Cross-references: UNIPARC:UPI00000B61C4; EMBL:U75366; NID:g2182812; PIDN:
R;Lewis, J.P.; Macrina, F.L.
Infect. Immun. 66, 3035-3042, 1998
A;Title: IS195, an insertion sequence-like element associated with protease genes in Por
A;Reference number: Z20844; MUID:98298016; PMID:9632563
A;Accession: T30526
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1350,'N',1352-1363,'Y',1365-1447,'H',1449-1732 <LEW>
A;Cross-references: UNIPARC:UPI00000BCA91; EMBL:AF017059; NID:g2738802; PID:g2738803; PI
R;Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
```

```
J. Biol. Chem. 269, 406-411, 1994
A;Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolati
A;Reference number: A53113; MUID:94103245; PMID:8276827
A;Accession: A53113
A;Status: preliminary
A;Molecule type: protein
A;Residues: 229-249 <PIK>
A;Cross-references: UNIPARC:UPI00000BA1A1
A;Experimental source: H66
A;Note: sequence extracted from NCBI backbone (NCBIP:141690)
C;Genetics:
A;Gene: prtP; prtK
C;Keywords: cysteine proteinase; hydrolase

Query Match 69.5%; Score 82; DB 2; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.00051;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NGGISLANYTGHGSETAW 19
||| ||| ||| ||| |||
Db 433 NTGVSFANYTAHGETAW 450

RESULT 5
S41406
arginine-specific cysteine proteinase - Porphyromonas gingivalis
C;Species: Porphyromonas gingivalis
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999
C;Accession: S41406
R;Sotiropoulos, C.; Kirsbaum, L.; Slakeski, N.; Jackson, C.; Reynolds, E.
submitted to the EMBL Data Library, January 1994
A;Description: Complete nucleotide sequence of the prtR gene encoding an arginine-specifi
A;Reference number: S41406
A;Accession: S41406
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <SOT>
A;Cross-references: UNIPARC:UPI000017AB4C; EMBL:Z29570; NID:g450920; PIDN:CAA82674.1; P

Query Match 44.9%; Score 53; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNGGISLANY 10
||||| |||||
Db 207 FNGGISLANY 216

RESULT 6
I39540
chitinase (EC 3.2.1.14) II precursor - Aeromonas sp.
C;Species: Aeromonas sp.
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C;Accession: I39540
R;Ueda, M.; Kawaguchi, T.; Arai, M.
J. Ferment. Bioeng. 78, 205-211, 1994
A;Title: Molecular cloning and nucleotide sequence of the gene encoding chitinase II fr
A;Reference number: I39540
A;Accession: I39540
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-542 <RES>
A;Cross-references: UNIPROT:Q59145; UNIPARC:UPI00000B97B1; GB:D31818; NID:g499711; PIDN:
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 41.9%; Score 49.5; DB 2; Length 542;
Best Local Similarity 42.9%; Pred. No. 13;
Matches 9; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNGGISLANYTGHGSETAWGT 21
||||| ||| ||| ||| |||
Db 401 YNGGATVA-YNGHNYQAKWWT 420
```

```
RESULT 7
Tl0727
C:protein kinase Xa21 (EC 2.7.1.1-) D, receptor type - long-staminate rice
C:Species: Oryza longistaminata (long-staminate rice)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: Tl0727
R;Song, W.Y.; Pi, L.Y.; Wang, G.L.; Gardner, J.; Holsten, T.; Ronald, P.C.
Plant Cell 9, 1279-1287, 1997
A:Title: Evolution of the rice Xa21 disease resistance gene family.
A:Reference number: Z15276; MUID:97432142; PMID:9286106
A:Accession: Tl0727
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-612 <SON>
A:Cross-references: UNIPROT:O24437; UNIPARC:UPI000009D7E7; EMBL:U72726; NID:G2586078; PID:
A:Experimental source: strain IRB821
C:Genetics:
C:Map position: 11
C:Keywords: phosphotransferase

Query Match 40.7%; Score 48; DB 2; Length 612;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

QY 1 FNGGISLA-NYTGHGSETAW 19
: || || | - || || |
Db 45 YQGQSLASWNTSGHGQCTW 65

RESULT 8
S68236
C:betaine/GABA transport protein BGT-1 - human
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S68236; I56522
R;Rasolla, A.; Galletta, L.J.V.; Barone, V.; Romeo, G.; Bagnasco, S.
FEBS Lett. 373, 229-233, 1995
A:Title: Molecular cloning and functional characterization of a GABA/betaine transporter
A:Reference number: S68236; MUID:96033979; PMID:7589472
A:Accession: S68236
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-614 <RAS>
A:Cross-references: UNIPROT:P48065; UNIPARC:UPI0000161F4B; EMBL:U27699; NID:G881474; PID:
R;Borden, L.A.; Smith, K.E.; Gustafson, E.L.; Brancheke, T.A.; Weinshank, R.L.
J. Neurochem. 64, 977-984, 1995
A:Title: Cloning and expression of a betaine/GABA transporter from human brain.
A:Reference number: I56522; MUID:95165166; PMID:7861179
A:Accession: I56522
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-9,'Y','ll-570,'QL',573-614 <RES>
A:Cross-references: UNIPARC:UPI00001354A9; GB:L423300; NID:G808695; PIDN:AAA65574.1; PID:
C:Genetics:
A:Gene: BGT-1
A:Map position: 12p13
C:Superfamily: gamma-aminobutyric acid transporter
C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
F;45-65/Domain: transmembrane #status predicted <TM1>
F;72-92/Domain: transmembrane #status predicted <TM2>
F;116-137/Domain: transmembrane #status predicted <TM3>
F;213-232/Domain: transmembrane #status predicted <TM4>
F;238-260/Domain: transmembrane #status predicted <TM5>
F;287-307/Domain: transmembrane #status predicted <TM6>
F;321-340/Domain: transmembrane #status predicted <TM7>
F;373-393/Domain: transmembrane #status predicted <TM8>
F;423-440/Domain: transmembrane #status predicted <TM9>
F;456-476/Domain: transmembrane #status predicted <TM10>
F;496-517/Domain: transmembrane #status predicted <TM11>
F;538-559/Domain: transmembrane #status predicted <TM12>
F;40,235/Binding site: phosphate (Thr) (covalent) #status predicted
F;171,183/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
F;418/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 40.7%; Score 48; DB 2; Length 614;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 ISLANYTGHGSETAW 19
: || || | - || || |
Db 94 VALQYTSQGSVTAW 108

RESULT 9
A41757
C:betaine transport protein, renal - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A41757
R;Yamauchi, A.; Uchida, S.; Kwon, H.M.; Preston, A.S.; Robey, R.B.; Garcia-Perez, A.; Bu
J. Biol. Chem. 267, 649-652, 1992
A:Title: Cloning of a Na(+)- and Cl(-)-dependent betaine transporter that is regulated b
A:Reference number: A41757; MUID:92112724; PMID:1370453
A:Accession: A41757
A:Molecule type: mRNA
A:Residues: 1-614 <YAM>
A:Cross-references: UNIPROT:P27799; UNIPARC:UPI00001354A8; GB:M80403; NID:G164031; PIDN:
C:Superfamily: gamma-aminobutyric acid transporter
C:Keywords: membrane protein

Query Match 40.7%; Score 48; DB 2; Length 614;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 ISLANYTGHGSETAW 19
: || || | - || || |
Db 94 VALQYTSQGSVTAW 108

RESULT 10
A27286
C:levanase (EC 3.2.1.65) precursor - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C:Accession: A27286; S06353; S11402; A69703
R;Schoegendorfer, K.; Schwab, H.; Lafferty, R.M.
Nucleic Acids Res. 15, 9606, 1987
A:Title: Nucleotide sequence of a cloned 2.5 kb PstI-EcoRI Bacillus subtilis DNA fragmen
A:Reference number: A27286; MUID:88067786; PMID:3120151
A:Accession: A27286
A:Molecule type: DNA
A:Residues: 1-677 <SCH>
A:Cross-references: UNIPROT:P05656; UNIPARC:UPI000016E952; GB:Y00485; NID:G40124; PIDN:
R;Martin, I.; Debarbouille, M.; Ferrari, E.; Klier, A.; Rapoport, G.
Mol. Gen. Genet. 208, 177-184, 1987
A:Title: Characterization of the levnanase gene of Bacillus subtilis which shows homology
A:Reference number: S06353; MUID:87286401; PMID:3112519
A:Accession: S06353
A:Molecule type: DNA
A:Residues: 1-657,'Q',659-677 <MAR>
A:Cross-references: UNIPARC:UPI00000607F3; EMBL:X05649
R;Martin-Verstraete, I.; Debarbouille, M.; Klier, A.; Rapoport, G.
J. Mol. Biol. 214, 657-671, 1990
A:Title: Levnanase operon of Bacillus subtilis includes a fructose-specific phosphotransf
A:Reference number: S11398; MUID:90355183; PMID:2117666
A:Accession: S11402
A:Molecule type: DNA
A:Residues: 1-10 <MA2>
A:Cross-references: UNIPARC:UPI000016E8BE; GB:X56098; NID:G39977; PIDN:CAA39581.1; PID:
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.
```

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69703
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-657; 'Q', 659-677 <KUN>
A;Cross-references: UNIPARC:UPI00000607F3; GB:299117; GB:AL009126; NID:g2634966; PIDN:CA
A;Experimental source: strain 168
A;Comment: The other genes in this fructose-inducible operon are levD, levE, levF, and I
C;Genetics:
A;Gene: sacC
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-677/Product: levansase #status predicted <MAT>

Query Match 40.7%; Score 48; DB 2; Length 677;
Best Local Similarity 40.0%; Pred. No. 28;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 2 NGGISLANYTGHGSETAWGT 21
|||: : : : :
Db 496 NGGVKVKSLTIHPLKKVWG 515

RESULT 11
T04313
protein kinase Xa21 (EC 2.7.1.-), receptor type - rice
C;Species: Oryza sativa (rice)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04313
R;Song, W.Y.; Pi, L.Y.; Wang, G.L.; Gardner, J.; Holsten, T.; Ronald, P.C.
Plant Cell 9, 1279-1287, 1997
A;Title: Evolution of the rice Xa21 disease resistance gene family.
A;Reference number: Z15276; MUID:97432142; PMID:9286106
A;Accession: T04313
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-813 <SON>
A;Cross-references: UNIPROT:O24435; UNIPARC:UPI00000AAEDA; EMBL:U72724; NID:g2586086; PI
A;Experimental source: cv. IRBB21
C;Genetics:
A;Map position: 11
C;Keywords: phosphotransferase

Query Match 40.7%; Score 48; DB 2; Length 813;
Best Local Similarity 47.6%; Pred. No. 34;
Matches 10; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

Qy 1 FNGGISLA--NYTGHGSETAW 19
: ||| ||| : ||| :
Db 47 YQGQSLASWNTSGHGQHCTW 67

RESULT 12
T10725
protein kinase Xa21 (EC 2.7.1.-) Al, receptor type - long-staminate rice
C;Species: Oryza longistaminata (long-staminate rice)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 31-Dec-2004
C;Accession: T10725
R;Song, W.Y.; Pi, L.Y.; Wang, G.L.; Gardner, J.; Holsten, T.; Ronald, P.C.
Plant Cell 9, 1279-1287, 1997
A;Title: Evolution of the rice Xa21 disease resistance gene family.
A;Reference number: Z15276; MUID:97432142; PMID:9286106
A;Accession: T10725
A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 1-996 <SON>
A;Cross-references: UNIPROT:O24436; UNIPARC:UPI00000A2B02; EMBL:U72725; NID:g2586079; PI
A;Experimental source: strain IRBB21
C;Genetics:
A;Map position: 11
A;Introns: 872/1
C;Superfamily: Receptor-like protein kinase; leucine-rich alpha-2-glycoprotein repeat hom
C;Keywords: phosphotransferase
F;77-100/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F;680-990/Domain: protein kinase homology <KIN>

Query Match 40.7%; Score 48; DB 2; Length 996;
Best Local Similarity 47.6%; Pred. No. 42;
Matches 10; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

Qy 1 FNGGISLA--NYTGHGSETAW 19
: ||| ||| : ||| :
Db 45 YQGQSLASWNTSGHGQHCTW 65

RESULT 13
A57676
protein kinase Xa21 (EC 2.7.1.-), receptor type precursor - rice
C;Species: Oryza sativa (rice)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 31-Dec-2004
C;Accession: A57676
R;Song, W.Y.; Wang, G.L.; Chen, L.L.; Kim, H.S.; Pi, L.Y.; Holsten, T.; Gardner, J.; Wan
Science 270, 1804-1806, 1995
A;Title: A receptor kinase-like protein encoded by the rice disease resistance gene, Xa2
A;Reference number: A57676; MUID:96106403; PMID:8525370
A;Accession: A57676
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1025 <SON>
A;Cross-references: UNIPROT:Q0640; UNIPARC:UPI00000A2DFF; GB:U37133; NID:g1122442; PIDN
C;Genetics:
A;Gene: Xa21
A;Note: confers resistance against the bacterial pathogen *Xanthomonas oryzae* pv. *oryzae*
C;Superfamily: Receptor-like protein kinase; leucine-rich alpha-2-glycoprotein repeat ho
C;Keywords: ATP; glycoprotein; magnesium; phosphotransferase; receptor; serine/threonine
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-1025/Product: protein kinase Xa21 #status predicted <MAT>
F;80-103/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;104-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;128-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;177-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;201-224/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;225-248/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;249-273/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;274-297/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;298-321/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F;322-345/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
F;352-375/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
F;377-400/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
F;401-424/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
F;425-448/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
F;449-472/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
F;473-496/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>
F;498-521/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>
F;522-545/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>
F;546-569/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR20>
F;570-593/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR21>
F;594-618/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR22>
F;706-1011/Domain: protein kinase homology <KIN>
F;714-722/Region: protein kinase ATP-binding motif
F;55_90_101_198_235_246_295_322_349_373_435_446_470_483_503_580_599/Binding site: carbo
F;736_752_842_844/Active site: Lys, Glu, Asp, Lys #status predicted
F;847_851/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 40.7%; Score 48; DB 1; Length 1025;
Best Local Similarity 47.6%; Pred. No. 44;

Matches 10; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

Qy 1 FNGGISLA--NYTGHGETAW 19
: ||||| : |||||
Db 45 YQGGQSLASWNTSGHGQCTW 65

RESULT 14

G83651
hypothetical protein BH0015 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: G83651
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-61 <STO>
A;Cross-references: UNIPROT:Q9KGP2; UNIPARC:UPI00000C377B; GB:AP001507; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0015

Query Match 39.8%; Score 47; DB 2; Length 61;
Best Local Similarity 52.9%; Pred. No. 2.9;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 3 GGISLANYTGHGETAW 19
|| : |||||
Db 42 GGRLLSTNWTGHPSEQHW 58

RESULT 15

AE2695
conserved hypothetical protein Atu0965 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AE2695
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthrner, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2695
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-294 <KUR>
A;Cross-references: UNIPROT:Q8UGS0; UNIPARC:UPI00000D19C5; GB:AE008688; PIDN:AAL41979.1;
C;Genetics:
A;Gene: Atu0965
A;Map position: circular chromosome

Query Match 39.8%; Score 47; DB 2; Length 294;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 GISLANYTGHGETA 18
| : | : |||||
Db 161 GVDLTRFTGHGSVAA 175

Search completed: August 25, 2006, 18:05:18
Job time : 27.4419 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 17:50:25 ; Search time 188.512 Seconds
(without alignments)
103.046 Million cell updates/sec

Title: US-10-387-977-1

Perfect score: 118

Sequence: 1 FNGGISLANYTHGHSETAMGT 21

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	118	100.0	736	1	CPG2_PORGI	P95493 porphyromon
2	118	100.0	1706	2	Q51838_PORGI	O51838 porphyromon
3	118	100.0	1706	2	O51839_PORGI	O51839 porphyromon
4	118	100.0	1706	2	O7MTE2_PORGI	O7MTE2 porphyromon
5	114	96.6	736	2	Q51844_PORGI	Q51844 porphyromon
6	114	96.6	991	1	CPGI_PORGI	P28784 porphyromon
7	114	96.6	1687	2	Q9R9B7_PORGI	O9R9B7 porphyromon
8	114	96.6	1704	2	Q51816_PORGI	O51816 porphyromon
9	100.5	85.2	422	2	Q51818_PORGI	O51818 porphyromon
10	82	69.5	1358	2	O6Q4T4_PORGI	O6Q4T4 porphyromon
11	82	69.5	1732	2	O07442_PORGI	O07442 porphyromon
12	82	69.5	1732	2	O52050_PORGI	O52050 porphyromon
13	82	69.5	1732	2	O51817_PORGI	O51817 porphyromon
14	75	63.6	1358	2	P96967_PORGI	P96967 porphyromon
15	75	63.6	1723	2	P72194_PORGI	P72194 porphyromon
16	75	63.6	1723	2	P72197_PORGI	P72197 porphyromon
17	60.5	51.3	459	2	Q4BEX8_BURVI	Q4BEX8 burkholderi
18	58.5	49.6	475	2	Q3RT78_RALME	Q3RT78 ralstonia m
19	57	48.3	559	2	Q2PJ92_ENTFA	Q2PJ92 enterococcu
20	57	48.3	559	2	O839T9_ENTFA	O839T9 enterococcu
21	55	46.6	527	2	O52PM7_9CAUD	O52PM7 xanthomonas
22	54.5	46.2	468	2	Q44Z96_9BURK	Q44Z96 burkholderi
23	54.5	46.2	468	2	Q4LK93_9BURK	Q4LK93 burkholderi
24	54	45.8	1123	2	O7MXX2_PORGI	O7MXX2 porphyromon
25	54	45.8	1357	2	Q2Z0E1_9BACT	Q2Z0E1 uncultured
26	53.5	45.3	468	2	Q39DM3_BURS3	Q39DM3 burkholderi
27	53	44.9	418	2	O7VD37_PROMA	O7VD37 prochloroco
28	52	44.1	157	2	Q9S8L6_STRTR	Q9S8L6 streptococc
29	52	44.1	357	2	Q3B616_PELUD	Q3B616 pelodictyon
30	52	44.1	408	2	O8Y1V4_RALSO	O8Y1V4 ralstonia s
31	51	43.2	156	2	Q47AY8_DECAR	Q47AY8 dechloromon

32	51	43.2	444	2	Q4H890_9DEIO	Q4H890 deinococcus
33	51	43.2	1090	2	Q9HJ11_THEAC	Q9HJ11 thermoplasm
34	51	43.2	1148	2	Q2PY26_9BACT	Q2PY26 uncultured
35	50.5	42.8	471	2	Q39LI4_BURS3	Q39LI4 burkholderi
36	50	42.4	434	2	Q35NH4_9BRAD	Q35NH4 bradyrhizob
37	50	42.4	491	2	Q8A8E5_BACTN	Q8A8E5 bacteroides
38	50	42.4	495	2	Q5LI05_BACFN	Q5LI05 bacteroides
39	50	42.4	526	2	Q6MS69_MYCMS	Q6MS69 mycoplasma
40	50	42.4	884	2	Q3E1V9_CHLAU	Q3E1V9 chloroflexu
41	50	42.4	926	2	Q4AIR5_9CHLB	Q4AIR5 chlorobium
42	50	42.4	1037	2	Q2R2D5_ORYSA	Q2R2D5 oryza sativ
43	50	42.4	3032	2	Q60XP7_CAEBR	Q60XP7 caenorhabdi
44	49.5	41.9	302	2	Q5TPA7_ANOGA	Q5TPA7 anopheles g
45	49.5	41.9	407	2	Q7Q1Q0_ANOGA	Q7Q1Q0 anopheles g

ALIGNMENTS

RESULT 1
CPG2_PORGI
ID_CPG2_PORGI STANDARD; PRT; 736 AA.
AC P95493; O33441;
DT 15-DEC-1998, integrated into UniProtKB/Swiss-Prot.
DT 03-OCT-2003, sequence version 2.
DT 07-MAR-2006, entry version 53.
DE Gingipain R2 precursor (EC 3.4.22.37) (Gingipain 2) (Arg-gingipain) (RGP-2).
DE (RGP-2).
GN Name=rgpB; Synonyms=prtRII, rgp2; OrderedLocusNames=PG0506;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PROTEIN SEQUENCE OF 230-651.
RC STRAIN=HG66;
RX MEDLINE=98370998; PubMed=9705298; DOI=10.1074/jbc.273.34.21648;
RA Potempa J., Mikolajczyk-Pawlinska J., Brassell D., Nelson D.,
RA Thøgersen I.B., Enghild J.J., Travis J.;
RT "Comparative properties of two cysteine proteinases (gingipains R),
RT the products of two related but individual genes of Porphyromonas
RT gingivalis.";
RL J. Biol. Chem. 273:21648-21657 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=W50;
RX MEDLINE=98304082; PubMed=9639929;
RA Slakeski N., Bhogal P.S., O'Brien-Simpson N.M., Reynolds E.C.;
RT "Characterization of a second cell-associated Arg-specific cysteine
RT proteinase of Porphyromonas gingivalis and identification of an
RT adhesin-binding motif involved in association of the prtR and prtK
RT proteinases and adhesins into large complexes.";
RL Microbiology 144:1583-1592 (1998).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601 (2003).
RN [4]
RP ENZYME REGULATION.
RX PubMed=11179305; DOI=10.1128/IAI.69.3.1402-1408.2001;
RA Gusman H., Travis J., Helmerhorst E.J., Potempa J., Troxler R.F.,
RA Oppenheim F.G.;
RT "Salivary histatin 5 is an inhibitor of both host and bacterial
RT enzymes implicated in periodontal disease.";

RL Infect. Immun. 69:1402-1408(2001).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=99452753; PubMed=10523290; DOI=10.1093/emboj/18.20.5453;
RA Eichinger A., Beisel H.-G., Jacob U., Huber R., Medrano F.-J.,
RA Banbula A., Potempa J., Travis J., Bode W.;
RT "Crystal structure of gingipain R: an Arg-specific bacterial cysteine
RT proteinase with a caspase-like fold.";
RL EMBL J. 18:5453-5462(1999).
CC -1- FUNCTION: Thiol protease which is believed to participate in
CC intracellular degradation and turnover of proteins. Its
CC proteolytic activity is a major factor in both periodontal tissue
CC destruction and in bacterial host defense mechanisms. Activates
CC complement C3 and C5 (By similarity).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins and small molecule
CC substrates, with a preference for Arg in P1.
CC -1- ENZYME REGULATION: Inhibited by human histatin-3 1/24 (histatin-
CC 5).
CC -1- SIMILARITY: Belongs to the peptidase C25 family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC -----
DR EMBL; U85038; AAB41892.1; -; Genomic DNA.
DR EMBL; AF007124; AAC26371.1; -; Genomic DNA.
DR EMBL; AE015924; AA065700.1; -; Genomic DNA.
DR PDB; 1CVR; X-ray; A=230-664.
DR MEROPS; C25.003; -.
DR GenomeReviews; AE015924_GR; PG0506.
DR TIGR; PG0506; -.
DR BioCyc; PGIN242619; PG0506-MONOMER; -.
DR InterPro; IPR001769; Peptidase_C25; -.
DR InterPro; IPR005536; Peptidase_C25_C.
DR InterPro; IPR012600; Propeptide_C25.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR Pfam; PF08126; Propeptide_C25; 1.
KW 3D-structure; Calcium; Complete proteome; Direct protein sequencing;
KW Hydrolase; Protease; Signal; Thiol protease; Virulence; Zymogen.
FT SIGNAL 1 24 By similarity.
FT PROPEP 25 229
FT CHAIN 230 736 /FTid=PRO_0000026535.
FT ACT SITE 440 440 Gingipain_R2.
FT ACT SITE 473 473 /FTid=PRO_0000026536.
FT CONFLICT 58 58 Proton donor.
FT CONFLICT 246 246 Nucleophile.
FT CONFLICT 251 251 G -> D (in Ref. 1).
FT CONFLICT 251 251 P -> A (in Ref. 1).
FT CONFLICT 254 254 E -> G (in Ref. 1).
FT CONFLICT 254 254 E -> K (in Ref. 1).
FT CONFLICT 398 398 I -> V (in Ref. 1).
FT CONFLICT 435 435 A -> V (in Ref. 1).
FT CONFLICT 480 482 YNV -> FSM (in Ref. 1).
FT CONFLICT 510 510 N -> D (in Ref. 1).
FT CONFLICT 512 512 S -> Y (in Ref. 1).
FT CONFLICT 515 515 S -> P (in Ref. 1).
FT CONFLICT 560 560 K -> N (in Ref. 1).
FT CONFLICT 582 582 K -> E (in Ref. 1).
FT TURN 237 238
FT STRAND 240 245
FT STRAND 247 252
FT HELIX 254 262
FT TURN 263 264
FT STRAND 266 271
FT STRAND 272 275
FT STRAND 277 278
FT HELIX 281 283
FT STRAND 294 294
FT TURN 295 296
FT STRAND 299 306
FT TURN 308 310
FT STRAND 311 311
FT STRAND 314 316

FT TURN 317 317
FT STRAND 319 322
FT HELIX 323 326
FT STRAND 329 333
FT STRAND 336 342
FT STRAND 345 346
FT HELIX 347 362
FT STRAND 363 363
FT TURN 366 367
FT TURN 369 372
FT STRAND 373 378
FT STRAND 380 380
FT STRAND 382 382
FT TURN 384 385
FT STRAND 386 386
FT STRAND 387 389
FT STRAND 391 391
FT HELIX 392 406
FT TURN 407 407
FT TURN 409 418
FT STRAND 421 430
FT HELIX 433 439
FT STRAND 441 441
FT STRAND 443 446
FT TURN 447 449
FT STRAND 452 452
FT TURN 453 455
FT HELIX 456 458
FT TURN 462 463
FT STRAND 467 474
FT TURN 475 476
FT TURN 478 479
FT STRAND 480 482
FT HELIX 485 491
FT STRAND 493 494
FT TURN 495 496
FT STRAND 497 498
FT STRAND 500 500
FT STRAND 502 509
FT STRAND 513 514
FT HELIX 515 527
FT TURN 528 529
FT STRAND 531 531
FT TURN 532 533
FT STRAND 538 538
FT HELIX 539 564
FT STRAND 565 569
FT TURN 571 572
FT STRAND 574 574
FT STRAND 578 578
FT STRAND 581 581
FT STRAND 585 586
FT STRAND 589 592
FT TURN 593 594
FT STRAND 596 604
FT TURN 605 606
FT STRAND 608 613
FT TURN 614 615
FT STRAND 616 622
FT STRAND 624 624
FT TURN 625 626
FT STRAND 627 631
FT STRAND 633 634
FT TURN 637 638
FT STRAND 640 647
FT TURN 649 650
FT STRAND 651 651
FT STRAND 654 661
SQ SEQUENCE 736 AA; 80967 MW; C848DD3FAB420833 CRC64;

Query Match 100.0%; Score 118; DB 1; Length 736;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7MTE2;
15-DEC-2003. integrated into UniProt

Q7MTE2;
15-DEC-2003. integrated into UniProt

```
DT 15-DEC-2003, sequence version 1.1.
DE 07-FEB-2006, entry version 12.
DE Hemagglutinin protein Hage.
GN Name=Hage; OrderedLocusNames=PG2024; ORFNames=PG_2024;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DE EMBL; AE015924; AA066991.1; -; Genomic_DNA.
DR HSP; P95493; 1CVR.
DR SMR; Q7MTE2; 228-655.
DR TIGR; PG2024; -.
DR BioCyc; PGIN242619:PG2024-MONOMER; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved_adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR InterPro; IPR012600; Propeptide_C25.
DR Pfam; PF07675; Cleaved_Adhesis_2.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25; 1.
DR Pfam; PF08126; Propeptide_C25; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1706 AA; 195673 MW; 6FE9B8AA9BA2815 CRC64;

Query Match 100.0%; Score 118; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
DB 426 FNGGISLANYTGHGSETAWGT 446

RESULT 5
Q51844 PORGI
ID Q51844 PORGI PRELIMINARY; PRT; 736 AA.
AC Q51844;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Arginine-specific cysteine proteinase (Arg-gingipain).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC33277;
RX MEDLINE=96007508; PubMed=7559528; DOI=10.1074/jbc.270.40.23619;
RA Nakayama K., Kadowaki T., Okamoto K., Yamamoto K.;
RT "Construction and characterization of arginine-specific cysteine

us-10-387-977-1.1.rup

RT proteinase (Arg-gingipain)-deficient mutants of Porphyromonas
RT gingivalis. Evidence for significant contribution of Arg-gingipain to
RT virulence.";
RL J. Biol. Chem. 270:23619-23626(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC33277;
RX MEDLINE=97276476; PubMed=9130229;
RA Nakayama K.;
RT "Domain-specific rearrangement between the two Arg-gingipain-encoding
RT genes in Porphyromonas gingivalis: possible involvement of
RT nonreciprocal recombination.";
RL Microbiol. Immunol. 41:185-196(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DE EMBL; D64081; BAA10963.1; -; Genomic_DNA.
DR HSP; P95493; 1CVR.
DR SMR; Q51844; 230-661.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR InterPro; IPR012600; Propeptide_C25.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25; 1.
DR Pfam; PF08126; Propeptide_C25; 1.
DR SEQUENCE 736 AA; 81004 MW; 5BB9C40E0BB4798 CRC64;

Query Match 96.6%; Score 114; DB 2; Length 736;
Best Local Similarity 95.2%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNGGISLANYTGHGSETAWGT 21
DB 428 FNGGISLANYTGHGSETAWGT 448

RESULT 6
CPGI_PORGI
ID CPGI_PORGI STANDARD; PRT; 991 AA.
AC P28784; Q45168;
DT 01-DEC-1992, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1997, sequence version 2.
DT 07-FEB-2006, entry version 47.
DE Gingipain R1 precursor (EC 3.4.22.37) (Gingipain 1) (Arg-gingipain)
DE (RGP-1).
GN Name=rgpA; Synonyms=rgpl;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PROTEIN SEQUENCE OF 228-290 AND
RP 517-541.
RC STRAIN=381;
RX MEDLINE=95168884; PubMed=7864651; DOI=10.1006/abbi.1995.1123;
RA Okamoto K., Misumi Y., Kadowaki T., Yoneda M., Yamamoto K.,
RA Ikehara Y.;
RT "Structural characterization of argingipain, a novel arginine-specific
RT cysteine proteinase as a major periodontal pathogenic factor from
RT Porphyromonas gingivalis.";
RL Arch. Biochem. Biophys. 316:917-925(1995).
RN [2]
RP PROTEIN SEQUENCE OF 228-270.
RC STRAIN=HG66;
RX MEDLINE=92406812; PubMed=1527017;
RA Chen Z., Potempa J., Polanowski A., Wikstrom M., Travis J.;
RT "Purification and characterization of a 50-kDa cysteine proteinase
RT (gingipain) from Porphyromonas gingivalis.";
RL J. Biol. Chem. 267:18896-18901(1992).
CC -!- FUNCTION: Thiol protease which is believed to participate in
```

DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved_adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25.C.
DR InterPro; IPR012600; Propeptide_C25.
DR Pfam; PF07675; Cleaved_Adhesin; 2.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25; 1.
DR Pfam; PF08126; Propeptide_C25; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1687 AA; 183703 MW; D085B516A399FE70 CR664;

Query Match 96.6%; Score 114; DB 2; Length 1687;
Best Local Similarity 95.2%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps

Qy 1 FNGGISLANYTGHGSETAWGT 21
||||| ||||| ||||| |||||
Db 407 FNGGISLVNVTGHGSETAWGT 427

RESULT 8
Q51816.PORGI PRELIMINARY; PRT; 1704 AA.
AC A51816;
DT 01-NOV-1996, integrated into UniProtKB/TREMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Arg-gingipain-1 proteinase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
[1]
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95138080; PubMed=7836351; DOI=10.1074/jbc.270.3.1007;
RA Pavloff N., Potempa J., Pike R.N., Prochazka V., Kiefer M.C.,
R Travis J., Barr P.J.;
RT "Molecular cloning and structural characterization of the Arg-gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a proteinase-adhesin polypeptide.";
RL J. Biol. Chem. 270:1007-1010(1995).

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>.
Distributed under the Creative Commons Attribution-NoDerivs License

ENBL; U15282; AAA69539.1; -; Genomic_DNA.
PIR; A55426; A55426.
HSP; P95493; ICVR.
SMR; Q51816; 228-635.
GO; GO:0008234; F:Cysteine-type peptidase activity; IEA.
GO; GO:0009405; P:pathogenesis; IEA.
GO; GO:0006508; P:proteolysis; IEA.
InterPro; IPR011628; Cleaved adhesin.
InterPro; IPR000977; DNA_ligase.
InterPro; IPR001769; Peptidase_C25.
InterPro; IPR005536; Peptidase_C25.C.
InterPro; IPR012600; Propeptide_C25.
Pfam; PF07675; Cleaved Adhesin; 2.
Pfam; PF01364; Peptidase_C25; 1.
Pfam; PF03785; Peptidase_C25; 1.
Pfam; PF08126; Propeptide_C25; 1.
PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
CHAIN 228 719 mature 50-kDa cysteine proteinase gingipain.

SQ SEQUENCE 1704 AA; 185437 MW; 6A34B40131C2A676 CR664;

Query Match 96.8%; Score 114; DB 2; Length 1704;
Best Local Similarity 95.2%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps

```

QY      1  FNGGISLVNYTGHGSETAWGT 21
DB      426 FNGGISLVNYTGHGSETAWGT 446

RESULT 9
Q51818 PORGI
ID Q51818 PORGI PRELIMINARY; PRT; 422 AA.
AC Q51818
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Cysteine proteinase alpha-gingipain.
GN Name: cgp8;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33277;
RA Garbia S.E., Shah H.N.;
RT "Molecular analysis of surface-associated enzymes of Porphyromonas
  gingivalis.";
RL Clin. Infect. Dis. 0:0-0(0).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33277;
RA Garbia S.E.;
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; X85186; CAA59469.1; -; Genomic_DNA.
DR HSSP; P95493; 1CVR.
DR SMR; O51818; 5-422.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR SEQUENCE 422 AA; 46280 MW; 773F18E28B4D4761 CRC64;

Query Match      85.2%; Score 100.5; DB 2; Length 422;
Best Local Similarity 86.4%; Pred. No. 7.9e-07;
Matches 19; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1  FNGGISLVNYTGHGSETAWGT 21
DB      202 FNGGISLVNYTGHGSDTAWGT 223

RESULT 10
Q604T4 PORGI
ID Q604T4 PORGI PRELIMINARY; PRT; 1358 AA.
AC Q604T4;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Lys-gingipain (Fragment).
GN Name: kgp;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W83variant;
RX PubMed=15297553; DOI=10.1128/JCM.42.8.3873-3876.2004;
RA Nadkarni M.A., Nguyen K.A., Chapple C.C., DeCarlo A.A., Jacques N.A.,

```

```

RA Hunter N.;
RT "Distribution of Porphyromonas gingivalis Biotypes Defined by Alleles
  of the kgp (Lys-Gingipain) Gene.";
RL J. Clin. Microbiol. 42:3873-3876(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY559244; AAS68176.1; -; Genomic_DNA.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved_adhesin.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 3.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
FT NON_TER 1
FT NCBI_TaxID=837;
SQ SEQUENCE 1358 AA; 146666 MW; 4B6A5D52729BFECB CRC64;

Query Match      69.5%; Score 82; DB 2; Length 1358;
Best Local Similarity 77.8%; Pred. No. 0.0023;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2  NGGISLVNYTGHGSETAW 19
DB      80 NTGVSFANYTAGSETAW 97

RESULT 11
O07442 PORGI
ID O07442 PORGI PRELIMINARY; PRT; 1732 AA.
AC O07442;
DT 01-JUL-1997, integrated into UniProtKB/TrEMBL.
DT 01-JUL-1997, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE Lysine-specific cysteine proteinase.
DE Name: prtK;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W50;
RX MEDLINE=99235907; PubMed=10219167;
RX DOI=10.1034/j.1399-302X.1999.140203.x;
RA Slakeski N., Cleal S.M., Bhogal P.S., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene prtK that encodes
  a lysine-specific cysteine proteinase and three sequence-related
  adhesins.";
RL Oral Microbiol. Immunol. 14:92-97(1999).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U75366; AAB60809.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved_adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 3.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
DR SEQUENCE 1732 AA; 187915 MW; 45D5B91377391703 CRC64;

```

```
Query Match      69.5%; Score 82; DB 2; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.003;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NGGISLANYTGHGSETAW 19
    | | | | | | | | | |
Db 433 NTGVSFANYTAHGETAW 450

RESULT 12
ID O52050_PORGI PRELIMINARY; PRT; 1732 AA.
AC O52050;
DT 01-JUN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JUN-1998, sequence version 1.
DE 07-FEB-2006, entry version 25.
DE Lysine specific cysteine protease.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W83;
RX MEDLINE=96298016; PubMed=9632563;
RA Lewis J.P., Macrina F.L.;
RT "TS195, an insertion sequence-like element associated with protease
  genes in Porphyromonas gingivalis.";
RL Infect. Immun. 66:3035-3042(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; AF017059; AAC26523.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR Pfam; PF01364; Cleaved Adhesin; 3.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1732 AA; 187932 MW; B2337463D5CB5EA5 CRC64;

Query Match      69.5%; Score 82; DB 2; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.003;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NGGISLANYTGHGSETAW 19
    | | | | | | | | | |
Db 433 NTGVSFANYTAHGETAW 450

RESULT 13
ID Q51817_PORGI PRELIMINARY; PRT; 1732 AA.
AC Q51817;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DE 07-FEB-2006, entry version 28.
DE Porphyropain.
GN Name=prtP;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RC STRAIN=W12;
RX MEDLINE=96213011; PubMed=8631659;
RA Barkocy-Gallagher G.A., Han N., Patti J.M., Whitlock J.,
RA Progulskie-Fox A., Lantz M.S.;
RT "Analysis of the prtP gene encoding porphyropain, a cysteine proteinase
  of Porphyromonas gingivalis.";
RL J. Bacteriol. 178:2734-2741(1996).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; U42210; AAB06565.1; -; Genomic_DNA.
DR PIR; T30836; T30836.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 3.
DR Pfam; PF01364; Peptidase_C25_C; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1732 AA; 187875 MW; 654271DBEF7BCAE4 CRC64;

Query Match      69.5%; Score 82; DB 2; Length 1732;
Best Local Similarity 77.8%; Pred. No. 0.003;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NGGISLANYTGHGSETAW 19
    | | | | | | | | | |
Db 433 NTGVSFANYTAHGETAW 450

RESULT 14
ID P36967_PORGI PRELIMINARY; PRT; 1358 AA.
AC P36967;
DT 01-MAY-1997, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1997, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE Hemagglutinin.
GN Name=hagD;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RA Han N., Lepine G., Whitlock J., Wojciechowski L., Progulskie-Fox A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; U68468; AAB49691.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 3.
DR Pfam; PF01364; Peptidase_C25_C; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1358 AA; 147102 MW; 47FCA0B25B06DED8 CRC64;
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 17:49:50 ; Search time 225.628 Seconds
(without alignments)
44.581 Million cell updates/sec

Title: US-10-387-977-2
Perfect score: 121
Sequence: 1 LNTGVSFANYTAHGETAWADP 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_8:*
1: geneseqp1990s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	121	100.0	22	2	AAW83071	Peptide f
2	121	100.0	506	2	AAW83126	AAW83126 PrtIIR48
3	121	100.0	1731	9	ADZ10439	P. gingiv
4	121	100.0	1731	9	ADZ10409	Adz10409 P. gingiv
5	121	100.0	1731	9	ADZ10393	P. gingiv
6	121	100.0	1731	9	ADZ10428	P. gingiv
7	121	100.0	1732	2	AAW96029	P. gingiv
8	121	100.0	1732	2	AAW24787	P. gingiv
9	121	100.0	1732	2	AAW69487	Haemagglu
10	114	94.2	509	9	ADV16839	Human cys
11	114	94.2	1358	2	AAW96032	P. gingiv
12	114	94.2	1358	2	AAW69494	Haemagglu
13	114	94.2	1723	9	ADV16837	Human cys
14	110	90.9	509	2	AAW34847	Lys-gingi
15	110	90.9	970	2	AAW72458	P. gingiv
16	82	67.8	21	2	AAW83070	Peptide f
17	82	67.8	493	2	AAW83125	P. gingiv
18	82	67.8	507	2	AAW83124	P. gingiv
19	82	67.8	1706	2	AAW24786	P. gingiv
20	82	67.8	1706	2	AAW24786	P. gingiv
21	82	67.8	1706	9	ADZ10438	P. gingiv
22	82	67.8	1706	9	ADZ10427	P. gingiv
23	82	67.8	1706	9	ADZ10456	P. gingiv

24	82	67.8	1706	9	ABW91466	Microbial
25	78	64.5	435	4	AAW49555	Gingipain
26	78	64.5	492	7	ABW02694	Porphym
27	78	64.5	492	9	ADV16838	Human cys
28	78	64.5	507	7	ABW02695	Porphym
29	78	64.5	737	2	AAW70186	Arg-gingi
30	78	64.5	737	2	AAW34846	Arg-gingi
31	78	64.5	737	3	AAW67395	Arg-gingi
32	78	64.5	737	4	AAU08937	P. gingiv
33	78	64.5	991	2	AAW77313	Porphym
34	78	64.5	1477	7	ABW02693	Porphym
35	78	64.5	1687	2	AAW96033	P. gingiv
36	78	64.5	1687	2	AAW69495	Haemagglu
37	78	64.5	1704	2	AAW70188	Arg-gingi
38	78	64.5	1704	2	AAW34843	Arg-gingi
39	78	64.5	1704	3	AAW67396	Arg-gingi
40	78	64.5	1704	4	AAU08938	P. gingiv
41	78	64.5	1704	9	ADV16836	Human cys
42	55	45.5	1123	9	ADZ10462	P. gingiv
43	55	45.5	1158	2	AAW34576	Porphym
44	55	45.5	1162	2	AAW34433	Porphym
45	50	41.3	450	8	ADN25795	Bacterial

ALIGNMENTS

RESULT 1
AAW83071
ID AAW83071 standard; peptide; 22 AA.
XX
AC AAW83071;
XX
XX
DT 03-MAR-1999 (first entry)
XX
DE Peptide from the 300kDa complex of cysteine proteinases and adhesins.
XX
XX
KW 300 kD multiprotein complex; cysteine proteinase; adhesin;
KW virulence factor; immune response; P. gingivalis infection;
KW periodontal disease; passive immunisation; subgingival plaque.
XX
OS Synthetic.
OS Porphyromonas gingivalis.
XX
XX
PN WO9849192-A1.
XX
PD 05-NOV-1998.
XX
PF 30-APR-1998; 98WO-AU000311.
XX
XX 30-APR-1997; 97AU-00006528.
XX
XX (UYME) UNIV MELBOURNE.
PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
PA (CSLC-) CSL LTD.
XX
PI Reynolds EC, Obrien-Simpson NM, Slakeski N;
XX WPI; 1999-024043/02.
DR
XX Immunogenic composition containing new Porphyromonas gingivalis peptides
PT - and related antibodies, useful for treatment, prevention and diagnosis
PT of periodontal disease.
XX
PS Claim 1; Page 49; 70pp; English.
XX
CC Peptides AAW83070-122 are derived from the 300 kD multiprotein complex of
CC cysteine proteinases and adhesins. The 300kDa multiprotein complex is the
CC major virulence factor of Porphyromonas gingivalis. The specification
CC describes a composition for raising an immune response against P.
CC gingivalis which contains at least one of the present peptides. The
CC products are used to treat and prevent P. gingivalis infection
CC (periodontal disease). Antibodies are used for passive immunisation, and

CC as diagnostic reagents for P. gingivalis in subgingival plaque. The
 CC peptides are used to detect P. gingivalis specific antibodies in serum

XX Sequence 22 AA;

Query Match 100.0%; Score 121; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 7.6e-12;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSPANTAHGSETAWADP 22
 |||||
 Db 1 LNTGVSPANTAHGSETAWADP 22
 |||||

RESULT 2

AAW83126

ID AAW83126 standard; protein; 506 AA.

XX AAW83126;

XX 03-MAR-1999 (first entry)

XX PrtIIK48 Lysine specific proteinase fragment.

XX 300 kD multiprotein complex; cysteine proteinase; adhesin;
 KW virulence factor; immune response; P. gingivalis infection;
 KW periodontal disease; passive immunisation; subgingival plaque.

XX Porphyromonas gingivalis.

XX WO9849192-A1.

XX 05-NOV-1998.

XX 30-APR-1998; 98WO-AU000311.

XX 30-APR-1997; 97AU-00006528.

XX (UTME) UNIV MELBOURNE.

PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.

PA (CSLC-) CSL LTD.

XX Reynolds EC, O'Brien-Simpson NM, Slakeski N;

XX WPI; 1999-024043/02.

XX Immunogenic composition containing new Porphyromonas gingivalis peptides
 PT - and related antibodies, useful for treatment, prevention and diagnosis
 PT of periodontal disease.

XX Example 1; Fig 8; 70pp; English.

XX The present sequence represents a fragment of the PrtIIK48 lysine
 CC specific proteinase of Porphyromonas gingivalis. The specification
 CC describes peptides derived from the 300 kD multiprotein complex of
 CC cysteine proteinases and adhesins. The 300kDa multiprotein complex is the
 CC major virulence factor of P. gingivalis. The specification describes a
 CC composition for raising an immune response against P. gingivalis which
 CC contains at least one of the present peptides. The products are used to
 CC treat and prevent P. gingivalis infection (periodontal disease).

XX Antibodies are used for passive immunisation, and as diagnostic reagents
 CC for P. gingivalis in subgingival plaque. The peptides are used to detect
 CC P. gingivalis specific antibodies in serum

XX Sequence 506 AA;

Query Match 100.0%; Score 121; DB 2; Length 506;

Best Local Similarity 100.0%; Pred. No. 3.4e-10;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSPANTAHGSETAWADP 22
 |||||

Db 204 LNTGVSPANTAHGSETAWADP 225
 |||||

RESULT 3

ID ADZ10439

XX ADZ10439 standard; protein; 1731 AA.

XX ADZ10439;

XX 16-JUN-2005 (first entry)

XX P. gingivalis lysine-specific cysteine protease SEQ ID 222.

XX Vaccine; enzyme; therapy; diagnosis; periodontal disease;

XX antiinflammatory; mouth disease; gingivitis; inflammation; mouth disease;

XX endocarditis; cardiac; cardiovascular disease; inflammation;

XX bacterial urinary tract infection; antibacterial; uropathic;

XX genitourinary disease; infection; osteomyelitis; osteopathic;

XX musculoskeletal disease; antigen.

XX Porphyromonas gingivalis.

XX WO2005019249-A2.

XX 03-MAR-2005.

XX 10-AUG-2004; 2004WO-US025778.

XX 15-AUG-2003; 2003US-0495589P.

XX (UYFL) UNIV FLORIDA.

XX Progulskke-Fox A, Hillman JD, Handfield M;

XX WPI; 2005-214225/22.

XX N-PSDB; ADZ10270.

XX New isolated immunogenic polypeptide, useful for diagnosing, treating,
 PT preventing, or ameliorating diseases and infections caused by
 PT Porphyromonas gingivalis e.g. localized prepubertal periodontitis.

XX Claim 1; SEQ ID NO 222; 73pp; English.

XX The invention relates to an isolated immunogenic Porphyromonas gingivalis
 CC polypeptide (I) comprising (a) a polypeptide at least about 5 contiguous
 CC amino acids of an amino acid sequence appearing as ADZ10384 - ADZ10443,
 CC ADZ10445 - ADZ10478 or ADZ10480-ADZ10571 (b) the polypeptide of (a) and
 CC a heterologous polypeptide or (c) an amino acid sequence selected from
 CC ADZ10384 - ADZ10443, ADZ10445 - ADZ10478 or ADZ10480-ADZ10571. Also
 CC included are an isolated polynucleotide (comprising a sequence that
 CC encodes (I) (a), (b) or (c)), an expression control sequence operably
 CC linked to the polynucleotide above (or the polynucleotide above and a
 CC heterologous polynucleotide), an expression vector comprising the
 CC polynucleotide, a host cell comprising the expression vector, an antibody
 CC (antibody fragment, or single-chain antibody) that specifically binds to
 CC the polypeptide, a composition comprising the antibody (antibody
 CC fragment, or single-chain antibody) and a pharmaceutical carrier,
 CC treating or preventing a disease/infection caused by Porphyromonas
 CC gingivalis, a composition comprising the isolated immunogenic polypeptide
 CC (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 CC response in an animal, detecting the presence of a first P. gingivalis
 CC polynucleotide in a test sample, detecting the presence of P.
 CC gingivalis antibody in a test sample, detecting the presence of P.
 CC gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting
 CC P. gingivalis infection in a subject and detecting P. gingivalis in a
 CC subject. The immunogenic polypeptide is useful for the diagnosis, caused
 CC treatment, prevention, and amelioration of diseases and infections caused
 CC by P. gingivalis. It is useful for treating or preventing localized
 CC prepubertal periodontitis, generalized prepubertal periodontitis,
 CC localized juvenile periodontitis, generalized juvenile periodontitis,
 CC rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 CC abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.

XX SQ Sequence 1731 AA;

Query Match 100.0%; Score 121; DB 9; Length 1731;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
 |||||

Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 4

ADZ10409

ID ADZ10409 standard; protein; 1731 AA.

XX AC ADZ10409;

DT 16-JUN-2005 (first entry)

XX DE P. gingivalis lysine-specific cysteine proteinase SEQ ID 192.

XX KW Vaccine; enzyme; therapy; diagnosis; periodontal disease;
 antiinflammatory; mouth disease; gingivitis; inflammation; mouth disease;
 endocarditis; cardiant; cardiovascular disease; inflammation;
 bacterial urinary tract infection; antibacterial; uropathic;
 genitourinary disease; infection; osteomyelitis; osteopathic;
 musculoskeletal disease; antigen.

XX OS Porphyromonas gingivalis.

XX PN WO2005019249-A2.

XX PD 03-MAR-2005.

XX PF 10-AUG-2004; 2004WO-US025778.

XX PR 15-AUG-2003; 2003US-0495589P.

XX PA (UYFL) UNIV FLORIDA.

XX PI Progulske-Fox A, Hillman JD, Handfield M;

XX DR WPI; 2005-214225/22.
 N-PSDB; ADZ10240.

XX PT New isolated immunogenic polypeptide, useful for diagnosing, treating,
 preventing, or ameliorating diseases and infections caused by
 Porphyromonas gingivalis e.g. localized prepubertal periodontitis.

XX PS Claim 1; SEQ ID NO 192; 73pp; English.

XX CC The invention relates to an isolated immunogenic Porphyromonas gingivalis
 polypeptide (I) comprising (a) a polypeptide at least about 5 contiguous
 amino acids of an amino acid sequence appearing as ADZ10384 - ADZ10443,
 ADZ10445 - ADZ10478 or ADZ10480-ADZ10571, (b) the polypeptide of (a) and
 a heterologous polypeptide or (c) an amino acid sequence selected from
 ADZ10384 - ADZ10443, ADZ10445 - ADZ10478 or ADZ10480-ADZ10571. Also
 included are an isolated polynucleotide (comprising a sequence that
 encodes (I) (a), (b) or (c)), an expression control sequence operably
 linked to the polynucleotide above (or the polynucleotide above and a
 heterologous polynucleotide), an expression vector comprising the
 polynucleotide, a host cell comprising the expression vector, an antibody
 (antibody fragment, or single-chain antibody) that specifically binds to
 the polypeptide, a composition comprising the antibody (antibody
 fragment, or single-chain antibody) and a pharmaceutical carrier,
 treating or preventing a disease/infection caused by Porphyromonas
 gingivalis, a composition comprising the isolated immunogenic polypeptide
 (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 response in an animal, detecting the presence of a first P. gingivalis
 polynucleotide in a test sample, detecting the presence of a P.
 gingivalis antibody in a test sample, detecting the presence of P.
 gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting

CC P. gingivalis infection in a subject and detecting P. gingivalis in a
 subject. The immunogenic polypeptide is useful for the diagnosis,
 treatment, prevention, and amelioration of diseases and infections caused
 by P. gingivalis . It is useful for treating or preventing localized
 prepubertal periodontitis, generalized prepubertal periodontitis,
 localized juvenile periodontitis, generalized juvenile periodontitis,
 rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.

XX SQ Sequence 1731 AA;

Query Match 100.0%; Score 121; DB 9; Length 1731;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
 |||||

Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 5

ADZ10393

ID ADZ10393 standard; protein; 1731 AA.

XX AC ADZ10393;

DT 16-JUN-2005 (first entry)

XX DE P. gingivalis lysine-specific cysteine proteinase SEQ ID 176.

XX KW Vaccine; enzyme; therapy; diagnosis; periodontal disease;
 antiinflammatory; mouth disease; gingivitis; inflammation; mouth disease;
 endocarditis; cardiant; cardiovascular disease; inflammation;
 bacterial urinary tract infection; antibacterial; uropathic;
 genitourinary disease; infection; osteomyelitis; osteopathic;
 musculoskeletal disease; antigen.

XX OS Porphyromonas gingivalis.

XX PN WO2005019249-A2.

XX PD 03-MAR-2005.

XX PF 10-AUG-2004; 2004WO-US025778.

XX PR 15-AUG-2003; 2003US-0495589P.

XX PA (UYFL) UNIV FLORIDA.

XX PI Progulske-Fox A, Hillman JD, Handfield M;

XX DR WPI; 2005-214225/22.
 N-PSDB; ADZ10228.

XX PT New isolated immunogenic polypeptide, useful for diagnosing, treating,
 preventing, or ameliorating diseases and infections caused by
 Porphyromonas gingivalis e.g. localized prepubertal periodontitis.

XX PS Claim 1; SEQ ID NO 176; 73pp; English.

XX CC The invention relates to an isolated immunogenic Porphyromonas gingivalis
 polypeptide (I) comprising (a) a polypeptide at least about 5 contiguous
 amino acids of an amino acid sequence appearing as ADZ10384 - ADZ10443,
 ADZ10445 - ADZ10478 or ADZ10480-ADZ10571, (b) the polypeptide of (a) and
 a heterologous polypeptide or (c) an amino acid sequence selected from
 ADZ10384 - ADZ10443, ADZ10445 - ADZ10478 or ADZ10480-ADZ10571. Also
 included are an isolated polynucleotide (comprising a sequence that
 encodes (I) (a), (b) or (c)), an expression control sequence operably
 linked to the polynucleotide above (or the polynucleotide above and a
 heterologous polynucleotide), an expression vector comprising the
 polynucleotide, a host cell comprising the expression vector, an antibody
 (antibody fragment, or single-chain antibody) that specifically binds to
 the polypeptide, a composition comprising the antibody (antibody
 fragment, or single-chain antibody) and a pharmaceutical carrier,
 treating or preventing a disease/infection caused by Porphyromonas
 gingivalis, a composition comprising the isolated immunogenic polypeptide
 (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 response in an animal, detecting the presence of a first P. gingivalis
 polynucleotide in a test sample, detecting the presence of a P.
 gingivalis antibody in a test sample, detecting the presence of P.
 gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting

CC (antibody fragment, or single-chain antibody) that specifically binds to
 CC the polypeptide, a composition comprising the antibody (antibody
 CC fragment, or single-chain antibody) and a pharmaceutical carrier,
 CC treating or preventing a disease/infection caused by Porphyromonas
 CC gingivalis, a composition comprising the isolated immunogenic polypeptide
 CC (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 CC response in an animal, detecting the presence of a first P. gingivalis
 CC polynucleotide in a test sample, detecting the presence of a P.
 CC gingivalis antibody in a test sample, detecting the presence of P.
 CC gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting
 CC P. gingivalis infection in a subject and detecting P. gingivalis in a
 CC subject. The immunogenic polypeptide is useful for the diagnosis,
 CC treatment, prevention, and amelioration of diseases and infections caused
 CC by P. gingivalis. It is useful for treating or preventing localized
 CC prepubertal periodontitis, generalized prepubertal periodontitis,
 CC localized juvenile periodontitis, generalized juvenile periodontitis,
 CC rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 CC abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.

XX SQ Sequence 1731 AA;

Query Match 100.0%; Score 121; DB 9; Length 1731;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LNTGVSFANYTAHGETAWADP 22
 |||||
 Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 6
 ADZ10428
 ID ADZ10428 standard; protein; 1731 AA.

AC ADZ10428;

DT 16-JUN-2005 (first entry)

DE P. gingivalis lysine-specific cysteine protease SEQ ID 211.

XX Vaccine; enzyme; therapy; diagnosis; periodontal disease;
 KW antiinflammatory; mouth disease; gingivitis; inflammation; mouth disease;
 KW endocarditis; cardiac; cardiovascular disease; inflammation;
 KW bacterial urinary tract infection; antibacterial; uropathic;
 KW genitourinary disease; infection; osteomyelitis; osteopathic;
 KW musculoskeletal disease; antigen.

XX Porphyromonas gingivalis.

OS WO2005019249-A2.

PN 03-MAR-2005.

XX 10-AUG-2004; 2004WO-US025778.

PF 15-AUG-2003; 2003US-0495589P.

PR (UYFL) UNIV FLORIDA.

XX Progulskes-Fox A, Hillman JD, Handfield M;

PI WPI: 2005-214225/22.

DR N-PSDB; ADZ10256.

XX New isolated immunogenic polypeptide, useful for diagnosing, treating,
 PT preventing, or ameliorating diseases and infections caused by
 PT Porphyromonas gingivalis e.g. localized prepubertal periodontitis.

XX Claim 1; SEQ ID NO 211; 73pp; English.

XX The invention relates to an isolated immunogenic Porphyromonas gingivalis

CC polypeptide (I) comprising (a) a polypeptide at least about 5 contiguous
 CC amino acids of an amino acid sequence appearing as ADZ10384 - ADZ10443,
 CC ADZ10445 - ADZ10478 or ADZ10480-ADZ10571, (b) the polypeptide of (a) and
 CC a heterologous polypeptide or (c) an amino acid sequence selected from
 CC ADZ10384 - ADZ10443, ADZ10445 - ADZ10478 or ADZ10480-ADZ10571. Also
 CC included are an isolated polynucleotide (comprising a sequence operably
 CC encoded (i) (a), (b) or (c)), an expression control sequence above and a
 CC heterologous polynucleotide), an expression vector comprising the
 CC polynucleotide, a host cell comprising the expression vector, an antibody
 CC (antibody fragment, or single-chain antibody) that specifically binds to
 CC the polypeptide, a composition comprising the antibody (antibody
 CC fragment, or single-chain antibody) and a pharmaceutical carrier,
 CC treating or preventing a disease/infection caused by Porphyromonas
 CC gingivalis, a composition comprising the isolated immunogenic polypeptide
 CC (or the polynucleotide) and a pharmaceutical carrier, eliciting an immune
 CC response in an animal, detecting the presence of a first P. gingivalis
 CC polynucleotide in a test sample, detecting the presence of a P.
 CC gingivalis antibody in a test sample, detecting the presence of P.
 CC gingivalis (or a P. gingivalis polypeptide) in a test sample, detecting
 CC P. gingivalis infection in a subject and detecting P. gingivalis in a
 CC subject. The immunogenic polypeptide is useful for the diagnosis,
 CC treatment, prevention, and amelioration of diseases and infections caused
 CC by P. gingivalis. It is useful for treating or preventing localized
 CC prepubertal periodontitis, generalized prepubertal periodontitis,
 CC localized juvenile periodontitis, generalized juvenile periodontitis,
 CC rapidly progressive adult periodontitis, refractory adult periodontitis,
 CC endocarditis, thyroid gland abscess, urinary tract infection, brain
 CC abscess, or vertebral osteomyelitis. The present sequence is a P.
 CC gingivalis immunogenic protein of the invention.

XX SQ Sequence 1731 AA;

Query Match 100.0%; Score 121; DB 9; Length 1731;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LNTGVSFANYTAHGETAWADP 22
 |||||
 Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 7

AAR96029

ID AAR96029 standard; protein; 1732 AA.

XX AAR96029;

DT 16-OCT-2003 (revised)

DT 04-SEP-1996 (first entry)

XX P. gingivalis porphyain.

XX Porphyain; haemagglutinin; periodontal disease; vaccine; antibody.

XX Porphyromonas gingivalis; strain W12.

XX Key Location/Qualifiers

FT Region 688..708
 FT /note= "Pro-Asn repeat region type 1"
 FT Region 887..952
 FT /note= "Pro-Asn repeat region type 2"
 FT Region 946..967
 FT /note= "Pro-Asn repeat region type 1"
 FT Region 985..1006
 FT /note= "Pro-Asn repeat region type 3"
 FT Region 1041..1100
 FT /note= "Pro-Asn repeat region type 4"
 FT Region 1341..1405
 FT /note= "Pro-Asn repeat region type 2"
 FT Region 1430..1451
 FT /note= "Pro-Asn repeat region type 3"
 FT Region 1488..1547

```

FT      Region      /note= "Pro-Asn repeat region type 4"
FT      1607. .1650
FT      /note= "Pro-Asn repeat region type 2"
XX      WO9617936-A2.
XX      13-JUN-1996.
XX      11-DEC-1995; 95WO-US016108.
XX      09-DEC-1994; 94US-00353485.
XX      (UYFL ) UNIV FLORIDA.
XX      (UABR-) UAB RES FOUND.
XX      Proguliske-Fox A, Tumwasorn S, Lepine G, Han N, Lantz M, Patti JM;
XX      WPI; 1996-287181/29.
XX      N-PSDB; AAT30653.
XX      Porphyromonas gingivalis genes and proteins - used in the detection and
XX      vaccination against periodontal disease.
XX      Claim 5; Page 76-81; 153pp; English.
XX      P. gingivalis W12 cysteine protease, porphyrain (AAR96029), was
XX      identified as the product of the prtP gene (AAT30653) isolated from P.
XX      gingivalis W12 genomic DNA. The porphyrain shows homology to the
XX      haemagglutinins (see also AAR96026-28 and AAR96030-33) of P. gingivalis
XX      318. It can be obtd. from transformed host cells and used as a vaccine to
XX      protect humans or animals against periodontal disease. Expression in
XX      Salmonella cells allows prodn. of a live vaccine. The porphyrain and
XX      haemagglutinins can also be used to detect the presence of anti-P.
XX      gingivalis antibodies and to raise monoclonal antibodies for diagnostic
XX      appln. (Updated on 16-OCT-2003 to standardise OS field)
XX      SQ      Sequence 1732 AA;

Query Match      100.0%; Score 121; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNTGVSFANYTAHGETAWADP 22
      |||||
Db      432 LNTGVSFANYTAHGETAWADP 453

RESULT 8
AAW24787
ID      AAW24787 standard; protein; 1732 AA.
AC      AAW24787;
XX
XX      17-OCT-2003 (revised)
DT      25-NOV-1997 (first entry)
XX
XX      PrtK antigenic protein complex.
DE
XX      Periodontal disease; cell surface protein; thiol protease; endopeptidase;
KW      PrtK; PrtK48; PrtK39; PrtK15; PrtK44; haemagglutinin; adhesin; therapy;
KW      diagnosis; vaccine; antigen.
XX
XX      Porphyromonas gingivalis; strain W50.
XX
XX      Key      Location/Qualifiers
FH      Peptide      1. .228
FT      /label= Pro-pro_peptide
FT      228. .229
FT      Cleavage-site
FT      Protein      229. .737
FT      /label= PrtK48
FT      /note= "48 kDa Lys-specific thiol protease"
FT      737. .738
FT      Cleavage-site
FT      Protein      738. .1156

/label= PrtK39
/note= "39 kDa adhesin"
FT      Cleavage-site
FT      Protein      1156. .1157
FT      /label= PrtK15
FT      /note= "15 kDa adhesin"
FT      Cleavage-site
FT      Protein      1291. .1292
FT      1292. .1732
FT      /label= PrtK44
FT      /note= "44 kDa adhesin"
XX      WO9716542-A1.
XX      09-MAY-1997.
XX      30-OCT-1996; 96WO-AU000673.
XX      30-OCT-1995; 95AU-00006275.
XX      (UYME ) UNIV MELBOURNE.
XX      (VICT-) VICTORIAN DAIRY IND AUTHORITY.
XX      Reynolds EC, Bhogal PS, Slakeski N;
XX      WPI; 1997-272112/24.
XX      N-PSDB; AAT78851.
XX      New antigenic protein complex from Porphyromonas gingivalis - comprising
XX      Arg- and Lys- specific thiol endo-peptidase(s), used in the detection,
XX      prevention and treatment of periodontal disease.
XX      Example 1; Fig 9b; 68pp; English.
XX      A PrtR-PrtK cell surface protein of Porphyromonas ginivalis (PG)
XX      comprises a 300 kDa complex composed a 48 kDa lysine-specific thiol
XX      protease and 39, 15 and 44 kDa adhesins encoded by the prtK gene
XX      (AAT78851), and a 45 kDa arginine-specific thiol protease and 44, 15, 17
XX      and 27 kDa adhesins (see AAW24786) encoded by the prtR gene (AAT78850). A
XX      claimed antigenic complex comprises at least one multimeric protein
XX      complex of PrtR and PrtK each containing at least one adhesin domain, the
XX      complex having a mol.wt. of over 200 kDa, and preferably comprises all 9
XX      proteins of the PrtR-PrtK complex (see also AAW24780-85). It can be used
XX      in a claimed composition to elicit an immune response directed against
XX      PG, and in a claimed method of reducing the prospect of PG infection
XX      and/or severity of disease. Antibodies directed against the complex are
XX      claimed for use in treating PG infection. Unlike whole PG cells or other
XX      previously prepared antigens based on fimbriae or the capsule, the PrtR-
XX      PrtK complex or component parts are safe and effective antigens. (Updated
XX      on 17-OCT-2003 to standardise OS field)
XX      SQ      Sequence 1732 AA;

Query Match      100.0%; Score 121; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNTGVSFANYTAHGETAWADP 22
      |||||
Db      432 LNTGVSFANYTAHGETAWADP 453

RESULT 9
AAW69487
ID      AAW69487 standard; protein; 1732 AA.
XX
XX      AAW69487;
XX
XX      22-DEC-1998 (first entry)
DT
XX      Haemagglutinin protein prtP.
DE
XX      Haemagglutinin protein; periodontal disease; vaccine; prtP.
KW
XX

```

```

OS Porphyromonas gingivalis.
XX US5824791-A.
XX 20-OCT-1998.
XX 11-DEC-1995; 95US-00570311.
XX PF 08-SEP-1988; 88US-00241640.
XX PR 25-JAN-1991; 91US-00647119.
XX PR 09-DEC-1994; 94US-00353485.
XX (UYFL ) UNIV FLORIDA.
XX (UABR-) UAB RES FOUND.
XX Patti JM, Han N, Lantz M, Tumwasorn S, Progulske-Fox A, Lepine G;
XX WPI; 1998-582627/49.
XX N-PSDB; AAV58874.
XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin and/or
XX protease poly:peptide(s) ).
XX Claim 1; Col 69-84; 101pp; English.
XX This sequence is encoded by a Porphyromonas gingivalis gene of the
XX invention. This sequence represents the prtP haemagglutinin protein. The
XX polypeptides are used to produce antibodies to organisms associated with
XX periodontal disease. The antibodies are also used in purification and
XX identification procedures. The genes and polypeptides are used as
XX vaccines against periodontal disease
XX SQ Sequence 1732 AA;

Query Match 100.0%; Score 121; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
|||||
Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 10
ADV16839
ID ADV16839 standard; protein; 509 AA.
XX AC ADV16839;
XX DT 24-FEB-2005 (first entry)
XX DE Human cysteine protease related protein, SEQ ID 4.
XX KW selectable marker; periodontal disease; angina; antiangiinal;
XX KW cardiovascular disease; diabetes; antidiabetic; cysteine protease;
XX KW enzyme.
XX OS Homo sapiens.
XX WO2004106541-A1.
XX 09-DEC-2004.
XX 31-MAY-2004; 2004WO-JP007867.
XX 30-MAY-2003; 2003JP-00154070.
XX (KYUS-) KYUSHU TLO CO LTD.
XX Yamamoto K;
XX WPI; 2005-021305/02.
XX

Query Match 94.2%; Score 114; DB 9; Length 509;
Best Local Similarity 90.9%; Pred. No. 4.4e-09;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
|||||
Db 204 LNTGVGFANYTAHGETSWADP 225

RESULT 11
AAR96032
ID AAR96032 standard; protein; 1358 AA.
XX AC AAR96032;
XX DT 16-OCT-2003 (revised)
XX DT 04-SEP-1996 (first entry)
XX DE P. gingivalis hagD haemagglutinin.
XX KW Haemagglutinin; hagD; periodontal disease; vaccine; antibody.
XX OS Porphyromonas gingivalis; strain FDC381.
XX WO9617936-A2.
XX 13-JUN-1996.
XX PF 11-DEC-1995; 95WO-US016108.
XX PR 09-DEC-1994; 94US-00353485.
XX (UYFL ) UNIV FLORIDA.
XX (UABR-) UAB RES FOUND.
XX Progulske-Fox A, Tumwasorn S, Lepine G, Han N, Lantz M, Patti JM;
XX WPI; 1996-287181/29.
XX N-PSDB; AAT30655.
XX Porphyromonas gingivalis genes and proteins - used in the detection and
XX vaccination against periodontal disease.
XX Claim 5; Page 125-129; 153pp; English.
XX P. gingivalis 381 haemagglutinin hagD (AAR96032) was identified as the
XX product of the second open reading frame of the hagD gene (AAR30655)
XX derived from P. gingivalis 318 genomic DNA. A first open reading frame
XX coded for hagD protease (see also AAR96031). The protease and
XX haemagglutinin can be obtd. from transformed host cells and used in
XX vaccines to protect humans or animals against periodontal disease.
XX Expression in Salmonella cells allows prodn. of live vaccines. The
XX haemagglutinin and protease can also be used to detect the presence of
XX anti-P. gingivalis antibodies and to raise monoclonal antibodies for

```

CC diagnostic appln. (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 1358 AA;

Query Match 94.2%; Score 114; DB 2; Length 1358;
Best Local Similarity 90.9%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||| ||||| ||||| |||||
Db 67 LNTGVGFANYTAHGETSWADP 88

RESULT 12
AAW69494
ID AAW69494 standard; protein; 1358 AA.

XX AC AAW69494;

XX DT 22-DEC-1998 (first entry)

XX DE Haemagglutinin protein hagD.

XX KW Haemagglutinin protein; periodontal disease; vaccine; hagD.

XX OS Porphyromonas gingivalis.

XX PN US824791-A.

XX PD 20-OCT-1998.

XX PF 11-DEC-1995; 95US-00570311.

XX PR 08-SEP-1988; 88US-00241640.

XX PR 25-JAN-1991; 91US-00647119.

XX PR 09-DEC-1994; 94US-00353485.

XX PA (UYFL) UNIV FLORIDA.

XX PA (UABR-) UAB RES FOUND.

XX PI Patti JM, Han N, Lantz M, Tunwasorn S, Progulske-Fox A, Lepine G;

XX DR N-PSDB; AAV58880.

XX PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin and/or

XX PS Claim 1; Col 145-158; 101pp; English.

XX CC This sequence is encoded by a Porphyromonas gingivalis gene of the
invention. This sequence represents the hagD haemagglutinin protein. The
polypeptides are used to produce antibodies to organisms associated with
periodontal disease. The antibodies are also used in purification and
identification procedures. The genes and polypeptides are used as
vaccines against periodontal disease

XX SQ Sequence 1358 AA;

Query Match 94.2%; Score 114; DB 2; Length 1358;
Best Local Similarity 90.9%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||| ||||| ||||| |||||
Db 67 LNTGVGFANYTAHGETSWADP 88

RESULT 13
ADV16837
ID ADV16837 standard; protein; 1723 AA.

XX AC ADV16837;

XX DT 24-FEB-2005 (first entry)
XX DE Human cysteine protease Arg-gingipain protein, rgpB.
XX KW selectable marker; periodontal disease; angina; antianginal;
XX KW cardiovascular disease; diabetes; antidiabetic; cysteine protease;
XX KW Arg-gingipain; rgp; enzyme; rgpB.
XX OS Homo sapiens.
XX PN WO2004106541-A1.
XX PD 09-DEC-2004.
XX PF 31-MAY-2004; 2004WO-JP007867.
XX PR 30-MAY-2003; 2003JP-00154070.
XX PA (KYUS-) KYUSHU TLO CO LTD.
XX PI Yamamoto K;
XX DR WPI; 2005-021305/02.
XX CC Novel periodontal disease marker for detecting periodontal disease,
comprising cysteine protease derived from Porphyromonas gingivalis,
useful for detecting periodontal disease, in saliva.
XX PS Claim 4; SEQ ID NO 2; 44pp; Japanese.
XX CC The invention relates to a novel periodontal disease marker for detecting
periodontal disease. The marker comprises a cysteine protease derived
from a Porphyromonas gingivalis microbe. The invention further comprises:
a periodontal disease ward appearance kit, comprising a sampling tool,
periodontal disease marker, synthetic substrate, buffer, and a reducing
agent. The periodontal disease marker is useful for detecting and
diagnosing periodontal disease from saliva. The periodontal disease
marker provides reliable and convenient detection of periodontal disease,
and provides simultaneous detection of angina and diabetes. This sequence
represents a cysteine protease human Arg-gingipain (rgp) protein of the
invention.
XX SQ Sequence 1723 AA;

Query Match 94.2%; Score 114; DB 9; Length 1723;
Best Local Similarity 90.9%; Pred. No. 1.9e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||| ||||| ||||| |||||
Db 432 LNTGVGFANYTAHGETSWADP 453

RESULT 14
AAW34847
ID AAW34847 standard; protein; 509 AA.

XX AC AAW34847;

XX DT 03-JUN-1998 (first entry)

XX DE Lys-gingipain high molecular weight prepolypeptide sequence.

XX KW Arg-specific gingipain protease; gingivalis; periodontal disease;
XX KW vaccine; infection.

XX OS Porphyromonas gingivalis.

XX PN WO9734629-A1.

XX PD 25-SEP-1997.

XX XX

```
PF 21-MAR-1997; 97WO-US004635.
XX
PR 22-MAR-1996; 96US-0013945P.
XX
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
PA (MORE-) MOREHOUSE SCHOOL MEDICINE.
XX
XX Potempa J, Travis J, Genco C;
XX
XX WPI; 1997-479993/44.
DR N-PSDB; AAT93874.
XX
XX Porphyromonas gingivalis Arg-specific gingipain protease peptide(s) -
PT useful for protecting animals and humans from gingivalis and periodontal
PT diseases.
XX
XX Disclosure; Page 76-78; 95pp; English.
XX
XX The present sequence represents a lysine-specific protease of
CC Porphyromonas gingivalis. The following peptides, derived from Arg- and
CC Lys-specific high molecular weight proteases, offer protection against
CC infection: YTYTVRDGK IKEGLTATTE DDGVTGNHE YCWEKXTAGS VSPKVC (I);
CC YTPVEEKQNG RMIVIVAKKY (II); QLPFIFDVAC VNGDFLFSMP CFAEALMRAQ (III);
CC GFENPYQPPVS NLTATTOGOK VTLKDAFSTK (IV); GNHEYCEVK YTAGVSPKVC KDVTV (V);
CC RMFMNYEPGR YTPVEEKQNG (VI); TFAGPEDTYK RMFMNYEPGR (VII); DYTITVYRDG
CC TKIKEGLTAT TPEEDGVTG NMEYCVCKY TAGVSPKVC (VIII); YTYTVRDGT KIKEGLTATTF
CC EDG (IX); RDGTKIKEGL TATTFEEDGV ATGN (X); KIKEGLTATT FEEDGVTGN HEY (XI)
CC ; KWDAPNGTPN PNPN PNPN PNPN FOTTLSE (XII); and YTPVEEKENG RMIVIVAKKY
CC (XIII). They are used in vaccines to protect animals, including humans,
CC from gingivitis and/or periodontal diseases
XX
XX Sequence 509 AA;
XX
Query Match 90.9%; Score 110; DB 2; Length 509;
Best Local Similarity 86.4%; Pred. No. 1.9e-08;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Oy 1 LNTGVSPFANYTAHGETAWADP 22
Db 204 LNTGVGFANYTVHGETSWADP 225
RESULT 15
AAR72458
ID AAR72458 standard; protein; 970 AA.
XX
AC AAR72458;
XX
DT 19-DEC-1995 (first entry)
XX
XX Porphyromonas gingivalis lysine-gingipain protein complex.
DE
XX Lysine-gingipain protein complex; amidolytic; proteolytic;
KW lysine specific proteinase; modulator identification; periodontitis;
KW therapy monitoring.
XX
XX OS Porphyromonas gingivalis.
XX
XX Key Location/Qualifiers
FH Peptide 1..228
FT Peptide /label= sig_peptide
FT Peptide 229..970
FT /label= mat_peptide
XX
XX WO9511298-A1.
XX
XX PN 27-APR-1995.
XX
XX PF 21-OCT-1994; 94WO-US012094.
XX
XX PR 21-OCT-1993; 93US-00141324.
XX
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
PA
```

```
XX Travis J, Potempa JS, Barr P, Pavloff N, Pike RN;
PI
XX WPI; 1995-170220/22.
DR N-PSDB; AAQ88141.
XX
XX Lys-gingipain complex prepn with amidolytic and proteolytic specificity -
PT for cleavage of an amide bond with Lysine contributing the carboxyl gp.
XX
XX Claim 2; Page 53-57; 75pp; English.
XX
XX AAQ88141 encodes AAR72458 the Porphyromonas gingivalis lysine-gingipain
CC protein complex (LGPC). The LGPC has amidolytic and proteolytic
CC specificity for an amide bond, where lysine contributes the carboxyl gp.,
CC i.e. a lysine specific protease. LGPC can be used to identify agents that
CC modulate the effect of LGPC on animals, and also for monitoring the
CC exposure of an animal to LGPC. Such a method can be used to monitor the
CC progress of a therapy designed to lessen the symptoms of periodontitis
XX
XX SQ Sequence 970 AA;
XX
Query Match 90.9%; Score 110; DB 2; Length 970;
Best Local Similarity 86.4%; Pred. No. 4.1e-08;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Oy 1 LNTGVSPFANYTAHGETAWADP 22
Db 432 LNTGVGFANYTVHGETSWADP 453
Search completed: August 25, 2006, 17:57:46
Job time : 231.628 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:04:46 ; Search time 36.3256 Seconds
(without alignments)
53.012 Million cell updates/sec

Title: US-10-387-977-2

Perfect score: 121

Sequence: 1 LNTGVSFANYAHGSETAWADP 22

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /EMC Celerra_SID33/ptodata/2/iaa/5 COMB.pep:*
2: /EMC Celerra_SID33/ptodata/2/iaa/6 COMB.pep:*
3: /EMC Celerra_SID33/ptodata/2/iaa/7 COMB.pep:*
4: /EMC Celerra_SID33/ptodata/2/iaa/H COMB.pep:*
5: /EMC Celerra_SID33/ptodata/2/iaa/PCTUS COMB.pep:*
6: /EMC Celerra_SID33/ptodata/2/iaa/RE COMB.pep:*
7: /EMC Celerra_SID33/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	121	100.0	1732	1	US-08-570-311-10 Sequence 10, Appl
2	121	100.0	1732	1	US-08-353-485-10 Sequence 10, Appl
3	121	100.0	1732	2	US-09-066-330-11 Sequence 11, Appl
4	114	94.2	942	1	US-08-141-324-14 Sequence 14, Appl
5	114	94.2	942	1	US-08-541-902-14 Sequence 14, Appl
6	114	94.2	1358	1	US-08-570-311-27 Sequence 27, Appl
7	110	90.9	509	2	US-08-822-324-8 Sequence 8, Appl
8	82	67.8	1706	2	US-09-066-330-10 Sequence 10, Appl
9	78	64.5	492	2	US-09-482-500A-2 Sequence 2, Appl
10	78	64.5	507	2	US-09-482-500A-3 Sequence 3, Appl
11	78	64.5	737	1	US-08-119-361-5 Sequence 5, Appl
12	78	64.5	737	2	US-08-336-308A-4 Sequence 4, Appl
13	78	64.5	737	2	US-08-822-324-4 Sequence 4, Appl
14	78	64.5	737	2	US-09-490-931-4 Sequence 4, Appl
15	78	64.5	1477	2	US-09-482-500A-1 Sequence 1, Appl
16	78	64.5	1687	1	US-08-570-311-29 Sequence 29, Appl
17	78	64.5	1704	2	US-08-336-308A-10 Sequence 10, Appl
18	78	64.5	1704	2	US-08-822-324-6 Sequence 6, Appl
19	78	64.5	1704	2	US-09-490-931-10 Sequence 10, Appl
20	46	38.0	321	2	US-10-152-886-21 Sequence 21, Appl
21	46	38.0	334	1	US-08-359-850-4 Sequence 4, Appl
22	46	38.0	355	1	US-07-946-497-5 Sequence 5, Appl
23	46	38.0	355	1	US-08-483-322-5 Sequence 5, Appl
24	46	38.0	355	1	US-08-478-882-5 Sequence 5, Appl
25	46	38.0	361	2	US-08-652-265-22 Sequence 22, Appl
26	46	38.0	361	2	US-08-834-497A-22 Sequence 22, Appl

ALIGNMENTS

RESULT 1

US-08-570-311-10
; Sequence 10, Application US/08570311
; Patent No. 5824791

GENERAL INFORMATION:

; APPLICANT: Progulsk-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph

; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ted W. Whitlock

; STREET: 2421 N.W. 41st Street, Suite A-1

; CITY: Gainesville

; STATE: FL

; COUNTRY: USA

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/570,311

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/353,485

; FILING DATE: 09-DEC-1994

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/647,119

; FILING DATE: 25-JAN-1991

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/241,640

; FILING DATE: 08-SEP-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Whitlock, Ted W.

; REGISTRATION NUMBER: 36,965

; REFERENCE/DOCKET NUMBER: UF15.C3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (904) 375-8100

; TELEFAX: (904) 372-5800

; INFORMATION FOR SEQ ID NO: 10:

Sequence 22, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 8, Appl
Sequence 12, Appl
Sequence 378, App
Sequence 2, Appl
Sequence 2, Appl
Sequence 8, Appl
Sequence 7681, Ap
Sequence 7615, Ap
Sequence 17, Appl
Sequence 251, App
Sequence 161, App

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1732 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-570-311-10

Query Match 100.0%; Score 121; DB 1; Length 1732;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSPFANYTAHGETAWADP 22
DB 432 LNTGVSPFANYTAHGETAWADP 453

RESULT 2
US-08-353-485-10
; Sequence 10, Application US/08353485
; Patent No. 5830710

; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph

; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C2
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1732 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-353-485-10

Query Match 100.0%; Score 121; DB 1; Length 1732;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSPFANYTAHGETAWADP 22

DB 432 LNTGVSPFANYTAHGETAWADP 453

RESULT 3

US-09-066-330-11
; Sequence 11, Application US/09066330A
; Patent No. 6511666
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: Bhogal, Peter S.

; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: Reynolds
; CURRENT APPLICATION NUMBER: US/09/066,330A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: EN 6275
; EARLIER FILING DATE: 1995-10-30
; EARLIER APPLICATION NUMBER: PCT/AU96/00673
; EARLIER FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 11
; LENGTH: 1732
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-09-066-330-11

Query Match 100.0%; Score 121; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSPFANYTAHGETAWADP 22
DB 432 LNTGVSPFANYTAHGETAWADP 453

RESULT 4

US-08-141-324-14
; Sequence 14, Application US/08141324
; Patent No. 5475097
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; APPLICANT: Pike, Robert N.

; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/141,324
; FILING DATE: 21-OCT-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 44-93
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-499-8080
; TELEFAX: 303-499-8089


```

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 942 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-141-324-14

Query Match          94.2%; Score 114; DB 1; Length 942;
Best Local Similarity 90.9%; Pred. No. 3.4e-09;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 432 LNTGVGFANYTAHGETSWADP 453

RESULT 5
US-08-541-902-14
; Sequence 14, Application US/08541902
; Patent No. 5707620
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; APPLICANT: Pike, Robert N.
; TITLE OF INVENTION: Lysine-specific Porphyromonas gingivalis
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/541,902
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/141,324
; FILING DATE: 21-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 44-93
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-499-8080
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 942 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-541-902-14

Query Match          94.2%; Score 114; DB 1; Length 942;
Best Local Similarity 90.9%; Pred. No. 3.4e-09;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 432 LNTGVGFANYTAHGETSWADP 453

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 942 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-141-324-14

Query Match          94.2%; Score 114; DB 1; Length 942;
Best Local Similarity 90.9%; Pred. No. 3.4e-09;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 432 LNTGVGFANYTAHGETSWADP 453

RESULT 6
US-08-570-311-27
; Sequence 27, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Progulsk-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1358 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-570-311-27

Query Match          94.2%; Score 114; DB 1; Length 1358;
Best Local Similarity 90.9%; Pred. No. 5.3e-09;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 67 LNTGVGFANYTAHGETSWADP 88

RESULT 7
US-08-822-324-8
; Sequence 8, Application US/08822324
; Patent No. 612917
; GENERAL INFORMATION:
; APPLICANT: Potempa, Jan S.
; APPLICANT: Travis, James
; APPLICANT: Genco, Caroline A.
```

```
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,945
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 103-95 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 488-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-822-324-8

Query Match 90.9%; Score 110; DB 2; Length 509;
Best Local Similarity 86.4%; Pred. No. 7e-09;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGSETAWDP 22
Db 204 LNTGVGFANYTVHGSETSWADP 225

RESULT 8
US-09-066-330-10
; Sequence 10, Application US/09066330A
; Patent No. 6511666
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: Bhogal, Peter S.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: Reynolds
; CURRENT APPLICATION NUMBER: US/09/066,330A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: PN 6275
; EARLIER FILING DATE: 1995-10-30
; EARLIER APPLICATION NUMBER: PCT/AU96/00673
; EARLIER FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-09-066-330-10

Query Match 67.8%; Score 82; DB 2; Length 1706;
Best Local Similarity 77.8%; Pred. No. 0.0007;
```

```
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NTGVSFANYTAHGSETAW 19
Db 427 NGGISLVNYTGHGSETAW 444

RESULT 9
US-09-482-500A-2
; Sequence 2, Application US/09482500A
; Patent No. 6627193
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Imamura, Takahisa
; APPLICANT: Potempa, Jan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION
; FILE REFERENCE: 235.00160101
; CURRENT APPLICATION NUMBER: US/09/482,500A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/115,869
; PRIOR FILING DATE: 1999-01-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-09-482-500A-2

Query Match 64.5%; Score 78; DB 2; Length 492;
Best Local Similarity 72.2%; Pred. No. 0.00067;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NTGVSFANYTAHGSETAW 19
Db 200 NGGISLVNYTGHGSETAW 217

RESULT 10
US-09-482-500A-3
; Sequence 3, Application US/09482500A
; Patent No. 6627193
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Imamura, Takahisa
; APPLICANT: Potempa, Jan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION
; FILE REFERENCE: 235.00160101
; CURRENT APPLICATION NUMBER: US/09/482,500A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/115,869
; PRIOR FILING DATE: 1999-01-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; US-09-482-500A-3

Query Match 64.5%; Score 78; DB 2; Length 507;
Best Local Similarity 72.2%; Pred. No. 0.0007;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NTGVSFANYTAHGSETAW 19
Db 200 NGGISLVNYTGHGSETAW 217

RESULT 11
US-08-119-361-5
; Sequence 5, Application US/08119361
; Patent No. 5523390
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis Arginine-specific Proteinase
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 10-SEP-1993
; APPLICATION NUMBER: US/08/119,361
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33878
; REFERENCE/DOCKET NUMBER: 21-93
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-499-8080
; TELEFAX: 303-499-8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-119-361-5

Query Match 64.5%; Score 78; DB 1; Length 737;
Best Local Similarity 72.2%; Pred. No. 0.0011;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTGVSFANYTAHGSETAW 19
| | | | | | | | | |
Db 427 NGGISLVNYTGHGSETAW 444

RESULT 12
US-08-336-308A-4
; Sequence 4, Application US/08336308A
; Patent No. 6017532
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis
; TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,308A
```

```
;
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/119,361
; FILING DATE: 10-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,441
; FILING DATE: 24-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 21-93C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-336-308A-4

Query Match 64.5%; Score 78; DB 2; Length 737;
Best Local Similarity 72.2%; Pred. No. 0.0011;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTGVSFANYTAHGSETAW 19
| | | | | | | | | |
Db 427 NGGISLVNYTGHGSETAW 444

RESULT 13
US-08-822-324-4
; Sequence 4, Application US/08822324
; Patent No. 6129917
; GENERAL INFORMATION:
; APPLICANT: Potempa, Jan S.
; APPLICANT: Travis, James
; APPLICANT: Genco, Caroline A.
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,945
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 103-95 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 488-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-822-324-4

Query Match 64.5%; Score 78; DB 2; Length 737;
Best Local Similarity 72.2%; Pred. No. 0.0011;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTGVSFANYTAHGSETAW 19
Db 427 NGGISLVNYTGHGSETAW 444

RESULT 14

US-09-490-931-4
; Sequence 4, Application US/09490931
; Patent No. 6274718

GENERAL INFORMATION:

; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S.
; APPLICANT: Barr, Philip J.
; APPLICANT: Pavloff, Nadine
; TITLE OF INVENTION: Porphyromonas gingivalis
; TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS: 16

; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,931
; FILING DATE:
; CLASSIFICATION:

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/336,308
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,441
; FILING DATE: 24-JUN-1994
; ATTORNEY/AGENT INFORMATION:

; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 21-93C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-490-931-4

Query Match 64.5%; Score 78; DB 2; Length 737;
Best Local Similarity 72.2%; Pred. No. 0.0011;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTGVSFANYTAHGSETAW 19
Db 427 NGGISLVNYTGHGSETAW 444

RESULT 15

US-09-482-500A-1
; Sequence 1, Application US/09482500A
; Patent No. 6627193
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Imamura, Takahisa
; APPLICANT: Potempa, Jan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BLOOD COAGULATION
; FILE REFERENCE: 235.00160101
; CURRENT APPLICATION NUMBER: US/09/482,500A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/115,869
; PRIOR FILING DATE: 1999-01-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1477
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-482-500A-1

Query Match 64.5%; Score 78; DB 2; Length 1477;
Best Local Similarity 72.2%; Pred. No. 0.0025;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTGVSFANYTAHGSETAW 19
Db 200 NGGISLVNYTGHGSETAW 217

Search completed: August 25, 2006, 18:06:28
Job time : 38.3256 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:21:12 ; Search time 126.884 Seconds
(without alignments)
80.315 Million cell updates/sec

Title: US-10-387-977-2

Perfect score: 121

Sequence: 1 LNTGVSPANTAHGSETAWADP 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /EMC_Celerra_SID33/ptodata/2/pubpaa/US07_PUBCOMB.psp.*
- 2: /EMC_Celerra_SID33/ptodata/2/pubpaa/US08_PUBCOMB.psp.*
- 3: /EMC_Celerra_SID33/ptodata/2/pubpaa/US09_PUBCOMB.psp.*
- 4: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10A_PUBCOMB.psp.*
- 5: /EMC_Celerra_SID33/ptodata/2/pubpaa/US10B_PUBCOMB.psp.*
- 6: /EMC_Celerra_SID33/ptodata/2/pubpaa/US11_PUBCOMB.psp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	22	4	US-10-387-977-2
2	121	100.0	509	4	US-10-387-977-101
3	121	100.0	1731	5	US-10-915-002-176
4	121	100.0	1731	5	US-10-915-002-192
5	121	100.0	1731	5	US-10-915-002-211
6	121	100.0	1731	5	US-10-915-002-222
7	121	100.0	1732	4	US-10-229-066-11
8	82	67.8	21	4	US-10-387-977-1
9	82	67.8	491	4	US-10-387-977-100
10	82	67.8	507	4	US-10-387-977-99
11	82	67.8	736	4	US-10-387-977-105
12	82	67.8	1706	4	US-10-229-066-10
13	82	67.8	1706	5	US-10-915-002-186
14	82	67.8	1706	5	US-10-915-002-210
15	82	67.8	1706	5	US-10-915-002-221
16	82	67.8	1706	5	US-10-915-002-239
17	82	67.8	1706	6	US-11-052-5548-176
18	55	45.5	1123	5	US-10-915-002-245
19	50	41.3	230	4	US-10-425-115-189940
20	50	41.3	450	4	US-10-369-493-8448
21	50	41.3	453	5	US-10-450-763-43923
22	50	41.3	475	6	US-11-087-099-5183
23	50	41.3	475	6	US-11-188-298-15792
24	50	41.3	878	4	US-10-437-963-155762
25	50	41.3	1528	4	US-10-437-963-146309
26	50	41.3	1843	4	US-10-437-963-117317
27	50	41.3	3041	4	US-10-437-963-202715

28	49	40.5	1534	4	US-10-437-963-189652	Sequence 189652,
29	48.5	40.1	731	4	US-10-437-963-189840	Sequence 189840,
30	48	39.7	444	5	US-10-481-265-9	Sequence 9, Appli
31	48	39.7	1686	4	US-10-437-963-181004	Sequence 181004,
32	47.5	39.3	445	4	US-10-081-872-116	Sequence 116, App
33	47.5	39.3	445	4	US-10-385-305-116	Sequence 116, App
34	47.5	39.3	463	4	US-10-081-872-318	Sequence 318, App
35	47.5	39.3	463	4	US-10-385-305-318	Sequence 318, App
36	47.5	39.3	464	4	US-10-081-872-126	Sequence 126, App
37	47.5	39.3	464	4	US-10-105-733-9	Sequence 9, Appli
38	47.5	39.3	464	4	US-10-081-739A-9	Sequence 9, Appli
39	47.5	39.3	464	4	US-10-385-305-126	Sequence 126, App
40	47.5	39.3	464	5	US-10-489-510-9	Sequence 9, Appli
41	47.5	39.3	471	4	US-10-081-872-112	Sequence 112, App
42	47.5	39.3	471	4	US-10-385-305-112	Sequence 112, App
43	47	38.8	321	4	US-10-424-569-239193	Sequence 239193,
44	47	38.8	359	5	US-10-450-763-54188	Sequence 54188, A
45	47	38.8	499	4	US-10-274-694-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-10-387-977-2
; Sequence 2, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-2

Query Match 100.0%; Score 121; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSPANTAHGSETAWADP 22

Db 1 LNTGVSPANTAHGSETAWADP 22

RESULT 2

US-10-387-977-101
; Sequence 101, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977

```
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-101

Query Match      100.0%; Score 121; DB 4; Length 509;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 204 LNTGVSFANYTAHGETAWADP 225

RESULT 3
US-10-915-002-176
; Sequence 176, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 176
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-176

Query Match      100.0%; Score 121; DB 5; Length 1731;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 4
US-10-915-002-192
; Sequence 192, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 192
; LENGTH: 1731
```

```
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-192

Query Match      100.0%; Score 121; DB 5; Length 1731;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 5
US-10-915-002-211
; Sequence 211, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 211
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-211

Query Match      100.0%; Score 121; DB 5; Length 1731;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 6
US-10-915-002-222
; Sequence 222, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 222
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-222

Query Match      100.0%; Score 121; DB 5; Length 1731;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTGVSFANYTAHGETAWADP 22
Db 432 LNTGVSFANYTAHGETAWADP 453
```

```
RESULT 7
US-10-229-066-11
; Sequence 11, Application US/10229066
; Publication No. US20030157637A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: Bhogal, Peter S.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: Reynolds
; CURRENT APPLICATION NUMBER: US/10/229,066
; CURRENT FILING DATE: 2002-08-28
; PRIOR FILING DATE: 1998-09-15
; PRIOR FILING DATE: 1998-09-15
; PRIOR FILING DATE: 1995-10-30
; PRIOR APPLICATION NUMBER: PCT/AU96/00673
; PRIOR FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1732
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-229-066-11

Query Match      100.0%; Score 121; DB 4; Length 1732;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LNTGVSPANYTAHGETAWADP 22
Db      432 LNTGVSPANYTAHGETAWADP 453
|||||
|||||

RESULT 8
US-10-387-977-1
; Sequence 1, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-1

Query Match      67.8%; Score 82; DB 4; Length 21;
Best Local Similarity 77.8%; Pred. No. 2.4e-05;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 NTGVSPANYTAHGETAW 19
Db      2 NGGISLANYTGHGETAW 19
|||
|||

RESULT 9
US-10-387-977-100
; Sequence 100, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-100

Query Match      67.8%; Score 82; DB 4; Length 491;
Best Local Similarity 77.8%; Pred. No. 0.00072;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 NTGVSPANYTAHGETAW 19
Db      200 NGGISLANYTGHGETAW 217
|||
|||

RESULT 10
US-10-387-977-99
; Sequence 99, Application US/10387977
; Publication No. US20040005276A1
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; FILE REFERENCE: 529282000301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-99

Query Match      67.8%; Score 82; DB 4; Length 507;
Best Local Similarity 77.8%; Pred. No. 0.00074;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 NTGVSPANYTAHGETAW 19
Db      200 NGGISLANYTGHGETAW 217
|||
|||
```



```
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 221
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-221

Query Match      67.8%; Score 82; DB 5; Length 1706;
Best Local Similarity 77.8%; Pred. No. 0.0028;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 NTGVSPFANYTAHGETAW 19
      |.:|:|:|:|:|:|:|:|:|
Db      427 NGGISLANYTGHGETAW 444

Search completed: August 25, 2006, 18:26:05
Job time : 129.884 secs
```

Result No.	Score	Query		Length	DB	ID	Description
		Match	Time				
1	48.5	40.1	731	6	US-10-449-902-54298	Sequence 54298, A	
2	48.5	40.1	731	6	US-10-449-902-54345	Sequence 54345, A	
3	48.5	40.1	731	6	US-10-449-902-54740	Sequence 54740, A	
4	46	38.0	361	7	US-11-327-689-22	Sequence 22, Appl	
5	46	38.0	373	6	US-10-449-902-48719	Sequence 48719, A	
6	46	38.0	614	7	US-11-300-928-62	Sequence 62, Appl	
7	45	37.2	317	7	US-11-056-3558-52484	Sequence 52484, A	
8	45	37.2	383	7	US-11-056-3558-52483	Sequence 52483, A	
9	45	37.2	392	6	US-10-449-902-29639	Sequence 29639, A	
10	45	37.2	392	6	US-10-449-902-49669	Sequence 49669, A	
11	45	37.2	435	7	US-11-056-3558-52482	Sequence 52482, A	
12	44	36.4	351	6	US-10-544-180-42	Sequence 42, Appl	
13	44	36.4	425	6	US-10-449-902-42939	Sequence 42939, A	
14	44	36.4	425	6	US-10-449-902-50133	Sequence 50133, A	
15	44	36.4	506	6	US-10-953-349-10213	Sequence 10213, A	
16	44	36.4	507	7	US-11-330-403-2356	Sequence 2356, Ap	
17	44	36.4	507	7	US-11-330-403-11763	Sequence 11763, A	
18	44	36.4	548	6	US-11-330-349-10212	Sequence 10212, A	
19	44	36.4	574	6	US-10-953-349-10211	Sequence 10211, A	
20	44	36.4	895	6	US-10-449-902-46345	Sequence 46345, A	
21	44	36.4	1101	6	US-10-449-902-44910	Sequence 44910, A	
22	43	35.5	22152	6	US-10-544-944-1	Sequence 1, Appli	
23	42.5	35.1	110	6	US-10-449-902-56113	Sequence 56113, A	
24	42.5	35.1	387	6	US-10-953-349-24666	Sequence 24666, A	
25	42.5	35.1	400	6	US-10-953-349-24665	Sequence 24665, A	

```

; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54345
; LENGTH: 731
; TYPE: prt
; ORGANISM: Oryza sativa
US-10-449-902-54345

```

Query Match 40.1%; Score 48.5; DB 6; Length 731;
Best Local Similarity 45.5%; Pred. No. 16;
Matches 10: Conservative 5; Mismatches 4; Indels 3; Gaps 1;

```

Qy      4  GVSFANYTAHGSET--AWADP 22
          |::| | | | |::| |
Db      16 GLSYDGYTDHGSQSDYFRFADP 37

```

```

RESULT 3
US-10-449-902-54740
; Sequence 54740, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54740
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-54740

```

Query Match	40.1%	Score 48.5;	DB 6;	Length 731;
Best Local Similarity	45.5%;	Pred. No. 16;		
Matches	10: Conservative	5: Mismatches	4: Indels	3: Gaps

QY 4 GVSFANYTAHGSET---AWADP 22
| : | : | | | : | :
db 16 GLSYDGYTDHGSODPYRFADP 37

RESULT 4
US-11-327-689-22
; Sequence 22, Application US/11327689
; Publication No. US2006017854A1
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; Drayna, Dennis T.
; Feder, John N.
; Gnirke, Andreas
; Ruddy, David
; Tsuchihashi, Zenta
; Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSQL for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/327,689
FILING DATE: 05-Jan-2006
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/497,957
FILING DATE: 04-FEB-2000
APPLICATION NUMBER: US/08/834,497
FILING DATE: 04-APR-1997
APPLICATION NUMBER: US 08/652,265
FILING DATE: 23-MAY-1996
APPLICATION NUMBER: US 08/632,673
FILING DATE: 16-APR-1996
APPLICATION NUMBER: US 08/630,912
FILING DATE: 04-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0056-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..361
OTHER INFORMATION: /note= "Rabbit leukocyte antigen (RLA)"
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-11-127-689-22


```
RESULT 10
US-10-449-902-49669
; Sequence 49669, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49669
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-49669

Query Match 37.2%; Score 45; DB 6; Length 392;
Best Local Similarity 43.8%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 7 FANYTAHGETAWADP 22
| : | : | : | : |
DB 122 FSELTSGATTWGGDP 137

RESULT 11
US-11-056-355B-52482
; Sequence 52482, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 52482
; LENGTH: 435
; TYPE: prt
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(435)
; OTHER INFORMATION: Ceres Seq. ID no. 15177286
US-11-056-355B-52482

Query Match 37.2%; Score 45; DB 7; Length 435;
Best Local Similarity 64.3%; Pred. No. 32;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 VSFANYTAHGETA 18
| : | : | : | : |
DB 267 VTFAGYWAYGSSTA 280

RESULT 12
US-10-544-180-42
; Sequence 42, Application US/10544180
; Publication No. US20060150280A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
```

```
; GENERAL INFORMATION:
; APPLICANT: Mohnen, Debra
; APPLICANT: Sterling, Jason D.
; APPLICANT: Doong, Ron L.
; APPLICANT: Koilli, Venkata S.K.
; APPLICANT: Hahn, Michael G.
; TITLE OF INVENTION: Galacturonosyltransferases, nucleic acids encoding same, and uses
; FILE REFERENCE: 14-03 US
; CURRENT APPLICATION NUMBER: US/10/544,180
; CURRENT FILING DATE: 2005-08-02
; PRIOR APPLICATION NUMBER: PCT/US2004/003545
; PRIOR FILING DATE: 2004-02-05
; PRIOR APPLICATION NUMBER: US 60/445,539
; PRIOR FILING DATE: 2003-02-06
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-544-180-42

Query Match 36.4%; Score 44; DB 6; Length 351;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 8 ANYTAHGETAWADP 22
| : | : | : | : |
DB 202 ANFTHYTRTFWSDP 216

RESULT 13
US-10-449-902-42939
; Sequence 42939, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42939
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-42939

Query Match 36.4%; Score 44; DB 6; Length 425;
Best Local Similarity 52.8%; Pred. No. 45;
Matches 10; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 2 NTGVSPFANYTAHGETAWA 20
| : | : | : | : |
DB 274 NGMTFAEYRAHFS--IWA 290

RESULT 14
US-10-449-902-50133
; Sequence 50133, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
```

; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 50133

; LENGTH: 425

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-449-902-50133

Query Match 36.4%; Score 44; DB 6; Length 425;
Best Local Similarity 52.6%; Pred.No. 45;
Matches 10; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 2 NTGVSPANYTAHGSETAWA 20

Db 274 NGGWTFAEYRAHFS--IWA 290

RESULT 15

US-10-953-349-10213

; Sequence 10213, Application US/109533349

; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THEREY

; FILE REFERENCE: 2750-1579PUS2

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 10213

; LENGTH: 506

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-10-953-349-10213

Query Match 36.4%; Score 44; DB 6; Length 506;
Best Local Similarity 52.9%; Pred.No. 54;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 GVSFANYTAHGSETAWA 20

Db 123 GLSMENRANGSEAWA 139

Search completed: August 25, 2006, 18:26:51

Job time : 22.9767 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:05:37 ; Search time 437.442 Seconds
(without alignments)
76.794 Million cell updates/sec

Title: US-10-387-977-2

Perfect score: 121

Sequence: 1 LNTGVSPANYTAHGETAWADP 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCTUS_COMB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066_COMB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073_COMB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/paa/US074_COMB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075_COMB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076_COMB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077_COMB.pep:*
8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078_COMB.pep:*
9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079_COMB.pep:*
10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080_COMB.pep:*
11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081_COMB.pep:*
12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082_COMB.pep:*
13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083_COMB.pep:*
14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084_COMB.pep:*
15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085_COMB.pep:*
16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086_COMB.pep:*
17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087_COMB.pep:*
18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088_COMB.pep:*
19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089_COMB.pep:*
20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090_COMB.pep:*
21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091_COMB.pep:*
22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092_COMB.pep:*
23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093_COMB.pep:*
24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094_COMB.pep:*
25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095_COMB.pep:*
26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096_COMB.pep:*
27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097_COMB.pep:*
28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098_COMB.pep:*
29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099_COMB.pep:*
30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100_COMB.pep:*
31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101_COMB.pep:*
32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102_COMB.pep:*
33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103_COMB.pep:*
34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104_COMB.pep:*
35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105_COMB.pep:*
36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106_COMB.pep:*
37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107_COMB.pep:*
38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108_COMB.pep:*
39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109_COMB.pep:*
40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110_COMB.pep:*
41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111_COMB.pep:*
42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112_COMB.pep:*
43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113_COMB.pep:*

44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114_COMB.pep:*
45: /EMC_Celerra_SIDS3/ptodata/2/paa/US600_COMB.pep:*
46: /EMC_Celerra_SIDS3/ptodata/2/paa/US601_COMB.pep:*
47: /EMC_Celerra_SIDS3/ptodata/2/paa/US602_COMB.pep:*
48: /EMC_Celerra_SIDS3/ptodata/2/paa/US603_COMB.pep:*
49: /EMC_Celerra_SIDS3/ptodata/2/paa/US604_COMB.pep:*
50: /EMC_Celerra_SIDS3/ptodata/2/paa/US605_COMB.pep:*
51: /EMC_Celerra_SIDS3/ptodata/2/paa/US606_COMB.pep:*
52: /EMC_Celerra_SIDS3/ptodata/2/paa/US607_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	22	24	US-09-423-056B-2
2	121	100.0	22	24	US-09-423-056B-2
3	121	100.0	22	33	US-10-387-977-2
4	121	100.0	509	24	US-09-423-056B-101
5	121	100.0	509	24	US-09-423-056B-101
6	121	100.0	509	33	US-10-387-977-101
7	121	100.0	1731	39	US-10-915-002-176
8	121	100.0	1731	39	US-10-915-002-192
9	121	100.0	1731	39	US-10-915-002-211
10	121	100.0	1731	39	US-10-915-002-222
11	121	100.0	1731	49	US-60-495-589-176
12	121	100.0	1731	49	US-60-495-589-192
13	121	100.0	1731	49	US-60-495-589-211
14	121	100.0	1731	49	US-60-495-589-222
15	121	100.0	1732	27	US-09-174-517-10
16	121	100.0	1732	27	US-09-791-537-18924
17	121	100.0	1732	32	US-10-229-066-11
18	114	94.2	1358	21	US-09-174-517-27
19	114	94.2	1723	27	US-09-791-537-86208
20	110	90.9	509	1	PCT-US97-04635-8
21	110	90.9	509	25	US-09-543-696B-8
22	110	90.9	509	25	US-09-543-696B-8
23	110	90.9	970	1	PCT-US94-12094-14
24	82	67.8	21	24	US-09-423-056B-1
25	82	67.8	21	24	US-09-423-056B-1
26	82	67.8	21	33	US-10-387-977-1
27	82	67.8	491	24	US-09-423-056B-100
28	82	67.8	491	24	US-09-423-056B-100
29	82	67.8	491	33	US-10-387-977-100
30	82	67.8	507	24	US-09-423-056B-99
31	82	67.8	507	33	US-10-387-977-99
32	82	67.8	507	33	US-09-423-056B-105
33	82	67.8	736	24	US-09-791-537-21158
34	82	67.8	736	33	US-10-387-977-105
35	82	67.8	1526	27	US-09-791-537-49353
36	82	67.8	1706	27	US-09-791-537-100403
37	82	67.8	1706	27	US-09-791-537-113656
38	82	67.8	1706	32	US-10-229-066-10
39	82	67.8	1706	39	US-10-915-002-186
40	82	67.8	1706	39	US-10-915-002-210
41	82	67.8	1706	39	US-10-915-002-221
42	82	67.8	1706	39	US-10-915-002-239
43	82	67.8	1706	49	US-11-052-554A-176
44	82	67.8	1706	49	US-60-495-589-186
45	82	67.8	1706	49	US-60-495-589-186

ALIGNMENTS

RESULT 1
US-09-423-056-2
; Sequence 2, Application US/09423056
; GENERAL INFORMATION:

us-10-387-977-2.rapm

Mon Aug 28 11:30:42 2006

```

; APPLICANT: Reynolds, Eric C.
; APPLICANT: O'Brien-Simpson, Neil M.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: Synthetic peptide constructs for the diagnosis and
; TITLE OF INVENTION: treatment of periodontitis associated with
; TITLE OF INVENTION: Porphyromonas gingivalis
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/423,056
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; ORGANISM: Porphyromonas gingivalis
US-09-423-056B-2

Query Match 100.0%; Score 121; DB 24; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGSETAWADP 22
Db 1 LNTGVSFANYTAHGSETAWADP 22

RESULT 2
US-09-423-056B-2
; Sequence 2 Application US/09423056B
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200300
; CURRENT APPLICATION NUMBER: US/09/423,056B
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; ORGANISM: Porphyromonas gingivalis
US-09-423-056B-2

Query Match 100.0%; Score 121; DB 24; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGSETAWADP 22
Db 1 LNTGVSFANYTAHGSETAWADP 22

RESULT 3
US-10-387-977-2
; Sequence 2 Application US/10387977
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200301

```

```

; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-2

Query Match 100.0%; Score 121; DB 33; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGSETAWADP 22
Db 1 LNTGVSFANYTAHGSETAWADP 22

RESULT 4
US-09-423-056-101
; Sequence 101 Application US/09423056
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: O'Brien-Simpson, Neil M.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: Synthetic peptide constructs for the diagnosis and
; TITLE OF INVENTION: treatment of periodontitis associated with
; TITLE OF INVENTION: Porphyromonas gingivalis
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/423,056
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
; ORGANISM: Porphyromonas gingivalis
US-09-423-056-101

Query Match 100.0%; Score 121; DB 24; Length 509;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGSETAWADP 22
Db 204 LNTGVSFANYTAHGSETAWADP 225

RESULT 5
US-09-423-056B-101
; Sequence 101 Application US/09423056B
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200300
; CURRENT APPLICATION NUMBER: US/09/423,056B
; CURRENT FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311

```



```
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-423-056B-101

Query Match      100.0%; Score 121; DB 24; Length 509;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSPFANYTAHGETAWADP 22
Db 204 LNTGVSPFANYTAHGETAWADP 225

RESULT 6
US-10-387-977-101
; Sequence 101, Application US/10387977
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric Charles
; APPLICANT: O'Brien-Simpson, Neil Martin
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS
; FILE REFERENCE: 52928200301
; CURRENT APPLICATION NUMBER: US/10/387,977
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 09/423,056
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00311
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: AU PO 6528
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-101

Query Match      100.0%; Score 121; DB 33; Length 509;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSPFANYTAHGETAWADP 22
Db 204 LNTGVSPFANYTAHGETAWADP 225

RESULT 7
US-10-915-002-176
; Sequence 176, Application US/10915002
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 176
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
```

```
US-10-915-002-176

Query Match      100.0%; Score 121; DB 39; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSPFANYTAHGETAWADP 22
Db 432 LNTGVSPFANYTAHGETAWADP 453

RESULT 8
US-10-915-002-192
; Sequence 192, Application US/10915002
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 192
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-192

Query Match      100.0%; Score 121; DB 39; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSPFANYTAHGETAWADP 22
Db 432 LNTGVSPFANYTAHGETAWADP 453

RESULT 9
US-10-915-002-211
; Sequence 211, Application US/10915002
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTIDE
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 211
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-211

Query Match      100.0%; Score 121; DB 39; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSPFANYTAHGETAWADP 22
Db 432 LNTGVSPFANYTAHGETAWADP 453

RESULT 10
US-10-915-002-222
```

```
; Sequence 222, Application US/10915002
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTI
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 222
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-222

Query Match      100.0%; Score 121; DB 39; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNTGVSFANYTAHGETAWADP 22
        |||
Db      432 LNTGVSFANYTAHGETAWADP 453

RESULT 11
US-60-495-589-176
; Sequence 176, Application US/60495589
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTI
; TITLE OF INVENTION: DIAGNOSIS, TREATMENT, AND MONITORING OF PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/60/495,589
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 176
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-60-495-589-176

Query Match      100.0%; Score 121; DB 49; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNTGVSFANYTAHGETAWADP 22
        |||
Db      432 LNTGVSFANYTAHGETAWADP 453

RESULT 12
US-60-495-589-192
; Sequence 192, Application US/60495589
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTI
; TITLE OF INVENTION: DIAGNOSIS, TREATMENT, AND MONITORING OF PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/60/495,589
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 192
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-387-977-2
```

```
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-60-495-589-192

Query Match      100.0%; Score 121; DB 49; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNTGVSFANYTAHGETAWADP 22
        |||
Db      432 LNTGVSFANYTAHGETAWADP 453

RESULT 13
US-60-495-589-211
; Sequence 211, Application US/60495589
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTI
; TITLE OF INVENTION: DIAGNOSIS, TREATMENT, AND MONITORING OF PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/60/495,589
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 211
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-60-495-589-211

Query Match      100.0%; Score 121; DB 49; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNTGVSFANYTAHGETAWADP 22
        |||
Db      432 LNTGVSFANYTAHGETAWADP 453

RESULT 14
US-60-495-589-222
; Sequence 222, Application US/60495589
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTI
; TITLE OF INVENTION: DIAGNOSIS, TREATMENT, AND MONITORING OF PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/60/495,589
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 222
; LENGTH: 1731
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-60-495-589-222

Query Match      100.0%; Score 121; DB 49; Length 1731;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNTGVSFANYTAHGETAWADP 22
        |||
Db      432 LNTGVSFANYTAHGETAWADP 453

RESULT 15
US-09-174-517-10
```

```
; Sequence 10, Application US/09174517
; GENERAL INFORMATION:
; APPLICANT: Progulske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patti, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/174,517
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; APPLICATION NUMBER: US 08/353,485
; FILING DATE: 09-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UFI5.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1732 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-174-517-10

Query Match 100.0%; Score 121; DB 21; Length 1732;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSPFANYTAHGSFADP 22
Db 432 LNTGVSPFANYTAHGSFADP 453

Search completed: August 25, 2006, 18:20:50
Job time : 439.442 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 18:06:46 ; Search time 27.6279 Seconds
(without alignments)
68.789 Million cell updates/sec

Title: US-10-387-977-2

Perfect score: 121

Sequence: 1 LNTGVSFANYTAHGETAWADP 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 363143 seqs, 86385820 residues

Total number of hits satisfying chosen parameters: 363143

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New *

- 1: /EMC_Celerra_SID33/prodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /EMC_Celerra_SID33/prodata/2/paa/US06_NEW_COMB.pep.*
- 3: /EMC_Celerra_SID33/prodata/2/paa/US07_NEW_COMB.pep.*
- 4: /EMC_Celerra_SID33/prodata/2/paa/US08_NEW_COMB.pep.*
- 5: /EMC_Celerra_SID33/prodata/2/paa/US09_NEW_COMB.pep.*
- 6: /EMC_Celerra_SID33/prodata/2/paa/US10_NEW_COMB.pep.*
- 7: /EMC_Celerra_SID33/prodata/2/paa/US11_NEW_COMB.pep.*
- 8: /EMC_Celerra_SID33/prodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	1732	6 US-10-229-066A-24	Sequence 24, Appl
2	82	67.8	1706	6 US-10-229-066A-21	Sequence 21, Appl
3	57	47.1	468	1 PCT-US06-18535-26266	Sequence 26266, A
4	57	47.1	468	7 US-11-431-855-26266	Sequence 26266, A
5	56	46.3	471	1 PCT-US06-18535-26339	Sequence 26339, A
6	56	46.3	471	7 US-11-431-855-26339	Sequence 26339, A
7	52	43.0	433	1 PCT-US06-18535-26437	Sequence 26437, A
8	52	43.0	433	7 US-11-431-855-26437	Sequence 26437, A
9	47.5	39.3	445	6 US-10-385-305A-116	Sequence 116, Appl
10	47.5	39.3	463	6 US-10-385-305A-318	Sequence 318, Appl
11	47.5	39.3	463	6 US-10-385-305A-323	Sequence 323, Appl
12	47.5	39.3	464	6 US-10-385-305A-126	Sequence 126, Appl
13	47.5	39.3	471	6 US-10-385-305A-112	Sequence 112, Appl
14	46.5	38.4	315	1 PCT-US06-18535-22801	Sequence 22801, A
15	46.5	38.4	315	7 US-11-431-855-22801	Sequence 22801, A
16	46	38.0	614	7 US-11-431-708-3588	Sequence 3588, Ap
17	46	38.0	614	7 US-11-431-708-3590	Sequence 3590, Ap
18	46	38.0	614	7 US-11-431-708-3601	Sequence 3601, Ap
19	46	38.0	614	7 US-11-431-708-3609	Sequence 3609, Ap
20	46	38.0	614	7 US-11-431-708-3614	Sequence 3614, Ap
21	46	38.0	614	7 US-11-475-062-6827	Sequence 6827, Ap
22	46	38.0	614	7 US-11-475-062-6829	Sequence 6829, Ap
23	46	38.0	614	7 US-11-475-062-6840	Sequence 6840, Ap
24	46	38.0	614	7 US-11-475-062-6848	Sequence 6848, Ap
25	46	38.0	614	7 US-11-475-062-6853	Sequence 6853, Ap

```

26 45 37.2 205 7 US-11-455-201-122 Sequence 122, App
27 45 37.2 205 7 US-11-478-144-778 Sequence 778, App
28 45 37.2 243 7 US-11-371-354-74171 Sequence 74171, A
29 44 36.4 2069 6 US-10-805-394A-4320 Sequence 4320, Ap
30 43.5 36.0 2544 1 PCT-US06-18535-23649 Sequence 23649, A
31 43.5 36.0 2544 7 US-11-431-855-23649 Sequence 23649, A
32 43 35.5 113 5 US-09-674-546B-976 Sequence 976, App
33 43 35.5 113 5 US-09-674-546B-980 Sequence 980, App
34 43 35.5 582 1 PCT-US06-30281-2649 Sequence 2649, Ap
35 43 35.5 582 1 PCT-US06-30281-20875 Sequence 20875, A
36 43 35.5 582 8 US-60-836-986-13044 Sequence 13044, A
37 43 35.5 1720 8 US-60-836-986-30809 Sequence 30809, A
38 43 35.5 14507 8 US-60-836-986-31920 Sequence 31920, A
39 42.5 35.1 3435 6 US-10-461-194A-132 Sequence 132, App
40 42 34.7 111 7 US-11-478-193-1051 Sequence 1051, Ap
41 42 34.7 116 1 PCT-US06-30281-6862 Sequence 6862, Ap
42 42 34.7 163 1 PCT-US06-18535-27585 Sequence 27585, A
43 42 34.7 163 7 US-11-431-855-27585 Sequence 27585, A
44 42 34.7 413 1 PCT-US06-18535-22856 Sequence 22856, A
45 42 34.7 413 7 US-11-431-855-22856 Sequence 22856, A

```

ALIGNMENTS

RESULT 1

```

US-10-229-066A-24
; Sequence 24, Application US/10229066A
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; APPLICANT: BHOGAL, PETER SINGH
; APPLICANT: SLAKESKI, NADA
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: 4137-20
; CURRENT APPLICATION NUMBER: US/10/229,066A
; CURRENT FILING DATE: 2002-08-28
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 09/066,330
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: PCT/AU96/00673
; PRIOR FILING DATE: 1996-10-30
; PRIOR APPLICATION NUMBER: AU PN 6275
; PRIOR FILING DATE: 1995-10-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 24
; LENGTH: 1732
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-229-066A-24

```

Query Match 100.0%; Score 121; DB 6; Length 1732;

Best Local Similarity 100.0%; Pred. No. 4.6e-10;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22

Db 432 LNTGVSFANYTAHGETAWADP 453

RESULT 2

```

US-10-229-066A-21
; Sequence 21, Application US/10229066A
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; APPLICANT: BHOGAL, PETER SINGH
; APPLICANT: SLAKESKI, NADA
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: 4137-20
; CURRENT APPLICATION NUMBER: US/10/229,066A
; CURRENT FILING DATE: 2002-08-28
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 09/066,330
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: PCT/AU96/00673

```

```
; PRIOR FILING DATE: 1996-10-30
; PRIOR APPLICATION NUMBER: AU PN 6275
; PRIOR FILING DATE: 1995-10-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 21
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-229-066A-21

Query Match      67.8%; Score 82; DB 6; Length 1706;
Best Local Similarity 77.8%; Pred. No. 0.00042;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 NTGVSFANYTAHGETAW 19
Db      427 NNGISLANYTGHGETAW 444

RESULT 3
PCT-US06-18535-26266
; Sequence 26266, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)A
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26266
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
PCT-US06-18535-26266

Query Match      47.1%; Score 57; DB 1; Length 468;
Best Local Similarity 83.3%; Pred. No. 0.86;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 TGVSFANYTAHG 14
Db      198 SGVGFANYTAHG 209

RESULT 4
US-11-431-855-26266
; Sequence 26266, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26266
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
US-11-431-855-26266

Query Match      47.1%; Score 57; DB 7; Length 468;
Best Local Similarity 83.3%; Pred. No. 0.86;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 TGVSFANYTAHG 14
Db      198 SGVGFANYTAHG 209

RESULT 5
US-11-431-855-26266
; Sequence 26266, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26266
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
US-11-431-855-26266

Query Match      47.1%; Score 57; DB 7; Length 468;
Best Local Similarity 83.3%; Pred. No. 0.86;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 TGVSFANYTAHG 14
Db      198 SGVGFANYTAHG 209

PCT-US06-18535-26339
; Sequence 26339, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)A
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26339
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
PCT-US06-18535-26339

Query Match      46.3%; Score 56; DB 1; Length 471;
Best Local Similarity 75.0%; Pred. No. 1.2;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 TGVSFANYTAHG 14
Db      188 SGIGFANYTAHG 199

RESULT 6
US-11-431-855-26339
; Sequence 26339, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26339
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R18194
US-11-431-855-26339

Query Match      46.3%; Score 56; DB 7; Length 471;
Best Local Similarity 75.0%; Pred. No. 1.2;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 TGVSFANYTAHG 14
Db      188 SGIGFANYTAHG 199

RESULT 7
PCT-US06-18535-26437
; Sequence 26437, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)A
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26437
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Burkholderia cepacia R1808
PCT-US06-18535-26437

Query Match      43.0%; Score 52; DB 1; Length 433;
Best Local Similarity 66.7%; Pred. No. 4.7;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 TGVSFANYTAHG 14
```

Db 163 SGIGFANYTGHG 174

RESULT 8

```

US-11-431-855-26437
; Sequence 26437, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26437
; LENGTH: 433
; TYPE: PR1
; ORGANISM: Burkholderia cepacia R1808
US-11-431-855-26437

```

Query Match 43.0%; Score 52; DB 7; Length 433;
Best Local Similarity 66.7%; Pred. No. 4.7;
Matches 8; Conservative 2; Mismatches 2; Indels

Qy	3	TGVSFANYTAHG	14
		: :	
Db	163	SGIGFANYTGHG	174

RESULT 9

```

US-10-385-305A-116
; Sequence 116, Application US/10385305A
; GENERAL INFORMATION:
; APPLICANT: CALLEN, Walter
; APPLICANT: RICHARDSON, Toby
; APPLICANT: FREY, Gerhard
; APPLICANT: GRAY, Kevin A.
; APPLICANT: KEROVUO, Janne S.
; APPLICANT: SLUPSKA, Malgorzata
; APPLICANT: BARTON, Nelson R.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; APPLICANT: SHORT, Jay M.
; TITLE OF INVENTION: AMYLASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462006120
; CURRENT APPLICATION NUMBER: US/10/385,305A
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/423,626
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 10/081,739
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/105,733
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 435
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 116
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Eukaryote
US-10-385-305A-116

```

Query Match 39.3%; Score 47.5; DB 6; Length 445;
Best Local Similarity 41.7%; Pred. No. 23;

	Matches	10;	Conservative	2;	Mismatches	9;	Indels	3;	Gaps	1;
Qy	1	LNTGVSPFANYTAH---	GSETAWAD	21						
Db	380	VYVGSKEFAGYTIHEYTGNLGGWID	403							

RESULT 10

```

US-10-385-305A-318
GENERAL INFORMATION:
GENERAL INFORMATION: Application US/10385305A
APPLICANT: CALLEN, Walter
APPLICANT: RICHARDSON, Toby
APPLICANT: GREY, Gerhard
APPLICANT: FREY, Kevin A.
APPLICANT: KEROVUO, Janne S.
APPLICANT: SLUPSKA, Malgorzata
APPLICANT: BARTON, Nelson R.
APPLICANT: O'DONOGHUE, Eileen
APPLICANT: MATHUR, Eric J.
APPLICANT: SHORT, Jay M.
TITLE OF INVENTION: AMPLASES, NUCLEIC
TITLE OF INVENTION: METHODS FOR MAKING
FILE REFERENCE: 564462006120
CURRENT APPLICATION NUMBER: US/10/385
CURRENT FILING DATE: 2003-03-06
PRIOR APPLICATION NUMBER: 10/081,872
PRIOR FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,495
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/270,496
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/291,122
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/423,626
PRIOR FILING DATE: 2002-10-31
PRIOR APPLICATION NUMBER: 10/081,739
PRIOR FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 10/105,733
PRIOR FILING DATE: 2002-03-22
NUMBER OF SEQ ID NOS: 435
SOFTWARE: FASTSEQ for Windows Version
SEQ ID NO 318
LENGTH: 463
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Clone
US-10-385-305A-318

```

Query Match	39.3%;	Score 47.5;	DB 6;	Length 463;
Best Local Similarity	41.7%;	Pred. No. 24;		
Matches 10;	Conservative	2;	Mismatches	9;
			Indels	3;
			Gaps	1;

QY 1 LNTGVSFANYTAH--GSETAWAD 21
:
DB 406 VNUGSKFAGYTIHEYTNLGGWVD 429

RESULT 11

US-10-385-305A-323
Sequence 323, Application US/10385305A
GENERAL INFORMATION:
APPLICANT: CALLEN, Walter
APPLICANT: RICHARDSON, Toby
APPLICANT: FREY, Gerhard
APPLICANT: GRAY, Kevin A.
APPLICANT: KEROUVO, Janne S.
APPLICANT: SLUPSKA, Malgorzata
APPLICANT: BARTON, Nelson R.
APPLICANT: O'DONOGHUE, Eileen
APPLICANT: MATHUR, Eric J.
APPLICANT: SHORT, Jay M.

```
; TITLE OF INVENTION: AMYLASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462006120
; CURRENT APPLICATION NUMBER: US/10/385,305A
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/423,626
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 10/081,739
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/105,733
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 435
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 323
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(27)
US-10-385-305A-323
```

```
Query Match 39.3%; Score 47.5; DB 6; Length 463;
Best Local Similarity 41.7%; Pred. No. 24;
Matches 10; Conservative 2; Mismatches 9; Indels 3; Gaps 1;
```

```
QY 1 LNTGVSPFANYTAH---GSETAWAD 21
Db 406 VNVGSKFAGYTIHEYTGNLGGWVD 429
```

```
RESULT 12
US-10-385-305A-126
; Sequence 126, Application US/10385305A
; GENERAL INFORMATION:
; APPLICANT: CALLEN, Walter
; APPLICANT: RICHARDSON, Toby
; APPLICANT: FREY, Gerhard
; APPLICANT: GRAY, Kevin A.
; APPLICANT: KEROVUO, Janne S.
; APPLICANT: SLUPSKA, Malgorzata
; APPLICANT: BARTON, Nelson R.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; APPLICANT: SHORT, Jay M.
; TITLE OF INVENTION: AMYLASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462006120
; CURRENT APPLICATION NUMBER: US/10/385,305A
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/423,626
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 10/081,739
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/105,733
```

```
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 435
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Environmental
US-10-385-305A-126
```

```
Query Match 39.3%; Score 47.5; DB 6; Length 464;
Best Local Similarity 41.7%; Pred. No. 24;
Matches 10; Conservative 2; Mismatches 9; Indels 3; Gaps 1;
```

```
QY 1 LNTGVSPFANYTAH---GSETAWAD 21
Db 407 VNVGSKFAGYTIHEYTGNLGGWVD 430
```

```
RESULT 13
US-10-385-305A-112
; Sequence 112, Application US/10385305A
; GENERAL INFORMATION:
; APPLICANT: CALLEN, Walter
; APPLICANT: RICHARDSON, Toby
; APPLICANT: FREY, Gerhard
; APPLICANT: GRAY, Kevin A.
; APPLICANT: KEROVUO, Janne S.
; APPLICANT: SLUPSKA, Malgorzata
; APPLICANT: BARTON, Nelson R.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; APPLICANT: SHORT, Jay M.
; TITLE OF INVENTION: AMYLASES, NUCLEIC ACIDS ENCODING THEM AND
; TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462006120
; CURRENT APPLICATION NUMBER: US/10/385,305A
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/423,626
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: 10/081,739
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/105,733
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 435
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Environmental
US-10-385-305A-112
```

```
Query Match 39.3%; Score 47.5; DB 6; Length 471;
Best Local Similarity 41.7%; Pred. No. 25;
Matches 10; Conservative 2; Mismatches 9; Indels 3; Gaps 1;
```

```
QY 1 LNTGVSPFANYTAH---GSETAWAD 21
Db 414 VNVGSKFAGYTIHEYTGNLGGWVD 437
```

```
RESULT 14
PCT-US06-18535-22801
; Sequence 22801, Application PC/TUS0618535
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
```

; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)A
; CURRENT APPLICATION NUMBER: PCT/US06/18535
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22801
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Magnaporthe grisea 70-15
PCT-US06-18535-22801

Query Match 38.4%; Score 46.5; DB 1; Length 315;
Best Local Similarity 40.9%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 4 GVSFANYTAHG---SETAWADP 22
Db 157 GARIVNITSHGHQLSDVWNSDP 178

RESULT 15
US-11-431-855-22801
; Sequence 22801, Application US/11431855
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et. al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53708)C
; CURRENT APPLICATION NUMBER: US/11/431,855
; CURRENT FILING DATE: 2006-05-10
; NUMBER OF SEQ ID NOS: 33637
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22801
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Magnaporthe grisea 70-15
US-11-431-855-22801

Query Match 38.4%; Score 46.5; DB 7; Length 315;
Best Local Similarity 40.9%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 4 GVSFANYTAHG---SETAWADP 22
Db 157 GARIVNITSHGHQLSDVWNSDP 178

Search completed: August 25, 2006, 18:21:50
Job time : 29.6279 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 17:58:43 ; Search time 24.5581 Seconds
(without alignments)
86.194 Million cell updates/sec

Title: US-10-387-977-2

Perfect score: 121

Sequence: 1 LNTGVSPANTAHGSETAWADP 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80.*

1: Piri.*

2: Piri2.*

3: Piri3.*

4: Piri4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	1732	2 T30836	lysine-specific cy
2	82	67.8	1526	2 S49763	gingipain R (EC 3
3	78	64.5	991	2 I40229	arginyln endopeptid
4	78	64.5	1704	2 A55426	gingipain R (EC 3
5	49.5	40.9	545	2 H86667	hypothetical prote
6	48	39.7	444	2 A70674	probable mbtC prot
7	47	38.8	307	2 E95939	probable inosine-u
8	47	38.8	311	2 G69731	PBSX prophage ORF
9	47	38.8	1672	2 C81675	polymorphic membra
10	46	38.0	197	2 F71248	probable proteasom
11	46	38.0	361	1 HLRB	MHC class I histoc
12	46	38.0	361	2 I46858	MHC class I RLA pr
13	46	38.0	503	2 B38745	cell adhesion mole
14	46	38.0	614	2 S68236	betaine/GABA trans
15	46	38.0	614	2 A41757	betaine transport
16	46	38.0	851	2 T31520	hypothetical prote
17	45.5	37.6	429	2 AH0630	4-hydroxyphenylace
18	45	37.2	214	2 C82950	glucose inhibited
19	45	37.2	219	2 H81107	uracil-DNA glycosy
20	45	37.2	219	2 B81908	probable uracil-DN
21	45	37.2	234	2 S37463	regulatory protein
22	45	37.2	348	2 D97490	hypothetical prote
23	45	37.2	379	2 E84274	membrane protein I
24	45	37.2	499	2 S46660	wetA protein - Pen
25	44	36.4	214	2 D89985	hypothetical prote
26	44	36.4	337	2 D87354	conserved hypothet
27	44	36.4	502	2 AH2289	4-alpha-glucanotra
28	44	36.4	506	2 AB2064	hypothetical prote
29	44	36.4	507	2 S52677	probable membrane

RESULT 1
T30836
lysine-specific cysteine proteinase porphypain (EC 3.4.22.-) - Porphyrromonas gingivalis
N:Alternate names: lysine-specific cysteine proteinase 1, 60K
C:Species: Porphyrromonas gingivalis
C:Date: 22-Oct-1999 #sequence revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30836; T30837; T30526; A53113
R:Barkocy-Gallagher, G.A.; Han, N.; Patti, J.M.; Whitlock, J.; Progulske-Fox, A.; Lantz
J.; Bacteriol. 178, 2734-2741, 1996
A:Title: Analysis of the prtp gene encoding porphypain, a cysteine proteinase of Porphy
A:Reference number: Z20895; MUID:96213011; PMID:8631659
A:Accession: T30836
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1732 <BAR>
A:Cross-references: UNIPROT:Q51817; UNIPARC:UPI00000893C0; EMBL:U42210; NID:gl314325; P
R:Slakeski, N.; Cleal, S.M.; Reynolds, E.C.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z20896
A:Accession: T30837
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-795, 'I', '797-1389, 'N', '1391-1478, 'Y', '1480-1732 <SLA>
A:Cross-references: UNIPARC:UPI00000861C4; EMBL:U75366; NID:g2182811; PID:g2182812; PID
R:Lewis, J.P.; Macrina, F.L.
Infect. Immun. 66, 3035-3042, 1998
A:Title: ISI95, an insertion sequence-like element associated with protease genes in Por
A:Reference number: Z20844; MUID:98298016; PMID:9632563
A:Accession: T30526
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1350, 'N', '1352-1363, 'Y', '1365-1447, 'H', '1449-1732 <LEW>
A:Cross-references: UNIPARC:UPI000008CA91; EMBL:AF017059; NID:g2738802; PID:g2738803; P
R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
J. Biol. Chem. 269, 406-411, 1994
A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolati
A:Reference number: A53113; MUID:94103245; PMID:8276827
A:Accession: A53113
A:Status: preliminary
A:Molecule type: protein
A:Residues: 229-249 <PIK>
A:Cross-references: UNIPARC:UPI000008A1A1
A:Experimental source: H66
A:Note: sequence extracted from NCBI backbone (NCBIP:141690)
C:Genetics:
A:Gene: prtp; prtK
C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 121; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

30 44 36.4 513 2 T11835
31 44 36.4 513 2 I37049
32 44 36.4 548 2 T49948
33 44 36.4 566 1 YDBPA7
34 44 36.4 1810 1 A32230
35 43.5 36.0 1490 2 S72351
36 43 35.5 119 2 PH0099
37 43 35.5 160 2 C81877
38 43 35.5 363 2 B45900
39 43 35.5 392 2 AF2072
40 43 35.5 448 2 T06698
41 43 35.5 463 1 T26378
42 43 35.5 505 2 S74648
43 43 35.5 511 1 B40164
44 43 35.5 523 2 F71952
45 43 35.5 523 2 D64555

ALIGNMENTS

QY 1 LNTGVSPFANYTAHGETAWADP 22
 |||||
 Db 432 LNTGVSPFANYTAHGETAWADP 453
 |||||
 RESULT 2
 S49763
 gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)
 C:Species: Porphyromonas gingivalis
 C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
 C:Accession: S49763
 R:Aduse-Opoku, J.; Muir, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A.
 submitted to the EMBL Data Library, November 1994
 A:Description: Cloning, sequence analysis and expression in Escherichia coli of prpR1 of
 A:Reference number: S49763
 A:Accession: S49763
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1526 <ADU>
 A:Cross-references: UNIPROT:Q51838; UNIPARC:UPI0000179912; EMBL:X82680
 C:Genetics:
 A:Gene: prpR1
 C:Keywords: cysteine proteinase; hydrolase
 Query Match 67.8%; Score 82; DB 2; Length 1526;
 Best Local Similarity 77.8%; Pred. No. 0.00032;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 NTGVSPFANYTAHGETAW 19
 |||||
 Db 427 NGGISLVNYTGHGETAW 444
 |||||
 RESULT 3
 I40229
 arginyl endopeptidase - Porphyromonas gingivalis
 C:Species: Porphyromonas gingivalis
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
 C:Accession: I40229
 R:Okamoto, K.; Misumi, Y.; Kadowaki, T.; Yoneda, M.; Yamamoto, K.; Ikehara, Y.
 Arch. Biochem. Biophys. 316, 917-925, 1995
 A:Title: Structural characterization of argingipain, a novel arginine-specific cysteine
 A:Reference number: I40229; MUID:95168884; PMID:7864651
 A:Accession: I40229
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-991 <RES>
 A:Cross-references: UNIPROT:P28784; UNIPARC:UPI000012829F; GB:D26470; NID:G927644; PIDN:
 Query Match 64.5%; Score 78; DB 2; Length 991;
 Best Local Similarity 72.2%; Pred. No. 0.00084;
 Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 NTGVSPFANYTAHGETAW 19
 |||||
 Db 427 NGGISLVNYTGHGETAW 444
 |||||
 RESULT 4
 A55426
 gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
 N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP; R
 C:Species: Porphyromonas gingivalis
 C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
 C:Accession: A55426; D53113
 R:Pavloff, N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr, B.
 J. Biol. Chem. 270, 1007-1010, 1995
 A:Title: Molecular cloning and structural characterization of the Arg-gingipain proteina
 A:Reference number: A55426; MUID:95138080; PMID:7836351
 A:Accession: A55426
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1704 <PAV>

A:Cross-references: UNIPROT:Q51816; UNIPARC:UPI000000B7BC1; GB:U15282; NID:G557067; PIDN:
 R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
 J. Biol. Chem. 269, 406-411, 1994
 A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolati
 A:Reference number: A53113; MUID:94103245; PMID:8276827
 A:Accession: D53113
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 228-249 <PIK>
 A:Cross-references: UNIPARC:UPI000000B9226
 A:Experimental source: H66
 A:Note: sequence extracted from NCBI backbone (NCBIP:141694)
 C:Keywords: cysteine proteinase; hydrolase
 Query Match 64.5%; Score 78; DB 2; Length 1704;
 Best Local Similarity 72.2%; Pred. No. 0.0015;
 Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 NTGVSPFANYTAHGETAW 19
 |||||
 Db 427 NGGISLVNYTGHGETAW 444
 |||||
 RESULT 5
 H86667
 hypothetical protein optA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
 C:Accession: H86667
 R:Boletin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: H86667
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-545 <STO>
 A:Cross-references: UNIPROT:Q9CIL2; UNIPARC:UPI000000C67F9; GB:AE005176; PID:G12723212; F
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: optA
 C:Superfamily: dipeptide transport protein
 Query Match 40.9%; Score 49.5; DB 2; Length 545;
 Best Local Similarity 42.3%; Pred. No. 11;
 Matches 11; Conservative 2; Mismatches 6; Indels 7; Gaps 1;
 QY 2 NTGVSPFANYTAH-----GSETAWA 20
 |||||
 Db 350 NTGVDFAKYAAQPKYDATAAKAAWA 375
 |||||
 RESULT 6
 A70674
 probable mbtC protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: A70674
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: A70674
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-444 <COL>
 A:Cross-references: UNIPROT:P71718; UNIPARC:UPI00001652E8; GB:Z81371; GB:AL123456; NID:
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: mbtC

R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux-Lecq, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koster, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Aethrich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier, lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F. Koster, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Aethrich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A:Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpst, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshihawa, H.F.; Zumbstein, E.; Yoshihikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580. PMID:98044033. PMID:9384377

F;418/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 38.0%; Score 46; DB 2; Length 614;
Best Local Similarity 53.3%; Pred. No. 43;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 VSFANYTAHGSETAW 19
|: ||: || ||
Db 94 VALGQYTSQGSVTAW 108

RESULT 15

A41757
betaine transport protein, renal - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A41757
R;Yamauchi, A.; Uchida, S.; Kwon, H.M.; Preston, A.S.; Robey, R.B.; Garcia-Perez, A.; Bu
J. Biol. Chem. 267, 649-652, 1992
A>Title: Cloning of a Na(+)- and Cl(-)-dependent betaine transporter that is regulated b
A;Reference number: A41757; MUID:92112724; PMID:1370453
A;Accession: A41757
A:Molecule type: mRNA
A;Residues: 1-614 <YAM>
A;Cross-references: UNIPROT:P27799; UNIPARC:UPI00001354A6; GB:M80403; NID:g164031; PIDN:
C:Superfamily: gamma-aminobutyric acid transporter
C;Keywords: membrane protein

Query Match 38.0%; Score 46; DB 2; Length 614;
Best Local Similarity 53.3%; Pred. No. 43;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 VSFANYTAHGSETAW 19
|: ||: || ||
Db 94 VALGQYTSQGSVTAW 108

Search completed: August 25, 2006, 18:05:14
Job time : 29.5581 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2006, 17:50:25 ; Search time 197.488 Seconds
(without alignments)
103.046 Million cell updates/sec

Title: US-10-387-977-2

Perfect score: 121

Sequence: 1 LNTGVSFANYTAHGETAWADP 22

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	1358	2	Q6Q4T4_PORGI
2	121	100.0	1732	2	O07442_PORGI
3	121	100.0	1732	2	O52050_PORGI
4	121	100.0	1732	2	O51817_PORGI
5	114	94.2	1358	2	P96967_PORGI
6	114	94.2	1723	2	P72194_PORGI
7	114	94.2	1723	2	P72197_PORGI
8	82	67.8	736	1	CPG2_PORGI
9	82	67.8	1706	2	O51838_PORGI
10	82	67.8	1706	2	O51839_PORGI
11	82	67.8	1706	2	Q7MRE2_PORGI
12	78	64.5	736	2	Q51844_PORGI
13	78	64.5	991	1	CPG1_PORGI
14	78	64.5	1687	2	Q9R9B7_PORGI
15	78	64.5	1704	2	O51816_PORGI
16	68.5	56.6	422	2	O51818_PORGI
17	57	47.1	468	2	Q39DM3_BURK
18	56	46.3	468	2	Q44296_9BURK
19	56	46.3	468	2	Q4LK93_9BURK
20	56	46.3	471	2	Q39L14_BURK
21	55	45.5	1123	2	Q7MXX2_PORGI
22	54	44.6	458	2	Q3JV75_BURP1
23	54	44.6	458	2	Q63W19_BURPS
24	53	43.8	459	2	Q2T0H9_BURTH
25	53	43.8	926	2	Q4AIR5_9CHLB
26	52	43.0	459	2	Q4BEX8_BURV1
27	51	42.1	527	2	Q12668_9BAST
28	51	42.1	847	2	Q90W12_ONCMY
29	51	42.1	893	2	Q6LFF5_PHOPR
30	50	41.3	297	2	Q3X551_9ACTN
31	50	41.3	359	2	Q89N78_BRAJA

32	50	41.3	459	2	Q3FLJ2_9BURK	O3flj2 burkholderi
33	50	41.3	475	2	Q3RT78_RALME	Q3rt78 ralstonia m
34	50	41.3	1175	2	Q5QJV8_9DIPT	Q5qjv8 culex trita
35	50	41.3	1473	2	Q7FA96_ORYSA	Q7fa96 oryza sativ
36	50	41.3	1528	2	Q7XKP3_ORYSA	Q7xkp3 oryza sativ
37	50	41.3	1535	2	Q53N07_ORYSA	Q53n07 oryza sativ
38	50	41.3	1735	2	Q6UU96_ORYSA	Q6uu96 oryza sativ
39	49.5	40.9	271	2	Q2JYD0_RHIET	Q2jydo rhizobium e
40	49.5	40.9	545	2	Q5CIL2_LACAU	Q5cil2 lactococcus
41	49	40.5	527	2	Q52PM7_9CAUD	Q52pm7 xanthomonas
42	49	40.5	798	2	Q6LL21_PHOPR	Q6ll21 photobacter
43	49	40.5	951	2	Q2X684_9GAMM	Q2x684 shewanella
44	49	40.5	951	2	Q2ZNF2_SHEPU	Q2zmf2 shewanella
45	49	40.5	1598	2	Q7XMF1_ORYSA	Q7xmfl oryza sativ

ALIGNMENTS

RESULT 1
ID Q6Q4T4_PORGI PRELIMINARY; PRT; 1358 AA.
AC Q6Q4T4;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Lys-gingipain (Fragment).
GN Name=kpg;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M83variant;
RX PubMed=15297553; DOI=10.1128/JCM.42.8.3873-3876.2004;
RA Nadkarni M.A., Nguyen K.A., Chapple C.C., DeCarlo A.A., Jacques N.A., Hunter N.;
RT "Distribution of Porphyromonas gingivalis Biotypes Defined by Alleles of the kpg (Lys-Gingipain) Gene.";
RL J. Clin. Microbiol. 42:3873-3876(2004).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

EMBL: AY559244; AAS68176.1; -: Genomic DNA.
DR GO: 0008234; F:cysteine-type peptidase activity; IEA.
DR GO: 0009405; P:pathogenesis; IEA.
DR GO: 0006508; P:proteolysis; IEA.
DR InterPro: IPR001769; Cleaved adhesin.
DR InterPro: IPR005536; Peptidase_C25.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF07675; Cleaved_Adhesin; 3_C.
DR Pfam: PF01364; Peptidase_C25; 1.
DR Pfam: PF03785; Peptidase_C25_C; 1.
FT NON_TER 1
FT NON_TER 1358
SQ SEQUENCE 1358 AA; 146666 MW; 4B6A5D52729BFECB CRC64;

Query Match 100.0%; Score 121; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
|||||
Db 79 LNTGVSFANYTAHGETAWADP 100

RESULT 2
O07442_PORGI PRELIMINARY; PRT; 1732 AA.
ID O07442_PORGI
AC O07442;
DT 01-JUL-1997, integrated into UniProtKB/TrEMBL.

DR	GO: 0009405; P: pathogenesis; IEA.
DR	GO: 0006508; P: proteolysis; IEA.
DR	InterPro: IPR011628; Cleaved_adhesin.
DR	InterPro: IPR000977; DNA_ligase.
DR	InterPro: IPR001769; Peptidase_C25.
DR	InterPro: IPR005536; Peptidase_C25_C.
DR	Pfam: PF07675; Cleaved_Adhesin_3.
DR	Pfam: PF01364; Peptidase_C35; 1.
DR	Pfam: PF03785; Peptidase_C25_C; 1.
DR	PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW	Protease.
SQL	SEQUENCE 1732 AA; 187932 MW; B2337463D5CB5EA5 CRC64;
Query Match 100.0%; Score 121; DB 2; Length 1732;	
Best Local Similarity 100.0%; Pred. No. 1.1e-09;	
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 LNTGVSFANYTAHGSETAWADP 22
DB	432 LNTGVSFANYTAHGSETAWADP 453
RESULT 4	
Q51817	PORGI
ID	Q51817_PORGI PRELIMINARY; PRT; 1732 AA.
AC	Q51817;
DT	01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT	01-NOV-1996, sequence version 1.
DT	07-FEB-2006, entry version 28.
DE	Porphyrapain.
GN	Name=prtp;
OS	Porphyromonas gingivalis (Bacteroides gingivalis).
OC	Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC	Porphyromonadaceae; Porphyromonas.
OX	NCBI_TaxID=837;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=W12;
RX	MEDLINE=96213011; PubMed=8631659;
RA	Barkocy-Gallagher G.A., Han N., Patti J.M., Whitlock J.,
RA	Prokulska-Fox A., Lantz M.S.;
RT	"Analysis of the prtp gene encoding porphyrapain, a cysteine proteinase
RT	of Porphyromonas gingivalis.";
RL	J. Bacteriol. 178:2734-2741(1996).
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	-----
EMBL	U42210; AAB06565.1; -; Genomic_DNA.
PIR	T30836; T30836.
DR	MEROPS; C25.002; -.
DR	GO: 0008234; F: cysteine-type peptidase activity; IEA.
DR	GO: 0009405; P: pathogenesis; IEA.
DR	GO: 0006508; P: proteolysis; IEA.
DR	InterPro: IPR011628; Cleaved_adhesin.
DR	InterPro: IPR000977; DNA_ligase.
DR	InterPro: IPR001769; Peptidase_C25.
DR	InterPro: IPR005536; Peptidase_C25_C.
DR	Pfam: PF07675; Cleaved_Adhesin_3.
DR	Pfam: PF01364; Peptidase_C25; 1.
DR	Pfam: PF03785; Peptidase_C25_C; 1.
DR	PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQL	SEQUENCE 1732 AA; 187875 MW; 654271DBEF7BCAA5 CRC64;
Query Match 100.0%; Score 121; DB 2; Length 1732;	
Best Local Similarity 100.0%; Pred. No. 1.1e-09;	
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 LNTGVSFANYTAHGSETAWADP 22
DB	432 LNTGVSFANYTAHGSETAWADP 453

```
RESULT 5
P96967_PORGI
ID P96967_PORGI PRELIMINARY; PRT; 1358 AA.
AC P96967;
DT 01-MAY-1997, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1997, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE Hemagglutinin.
GN Name=hagD;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RA Han N., Lepine G., Whitlock J., Wojciechowski L., Progulskie-Fox A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U68468; BAB49691.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 3.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
SQ SEQUENCE 1358 AA; 147102 MW; 47FCA0B25B06DED8 CRC64;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RA Han N., Lepine G., Whitlock J., Wojciechowski L., Progulskie-Fox A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U68468; BAB49691.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 3.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
SQ SEQUENCE 1358 AA; 147102 MW; 47FCA0B25B06DED8 CRC64;

Query Match 94.2%; Score 114; DB 2; Length 1358;
Best Local Similarity 90.9%; Pred. No. 1.1e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||||| |||||
Db 67 LNTGVGFANYTAHGETSWADP 88

RESULT 6
P72194_PORGI
ID P72194_PORGI PRELIMINARY; PRT; 1723 AA.
AC P72194;
DT 01-FEB-1997, integrated into UniProtKB/TrEMBL.
DT 01-FEB-1997, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Lys-gingipain.
GN Name=kgp;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RA Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT "Cloning and sequencing of the gene encoding a novel lysine-specific
RT cysteine proteinase (lys-gingipain) in Porphyromonas gingivalis:
RT structural relationship with the arginine-specific cysteine proteinase
RT (Arg-gingipain).";
RL J. Biochem. 120:398-406 (1996).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

Qy 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||||| |||||
Db 67 LNTGVGFANYTAHGETSWADP 88

Query Match 94.2%; Score 114; DB 2; Length 1358;
Best Local Similarity 90.9%; Pred. No. 1.1e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||||| |||||
Db 67 LNTGVGFANYTAHGETSWADP 88

RESULT 7
P72197_PORGI
ID P72197_PORGI PRELIMINARY; PRT; 1723 AA.
AC P72197;
DT 01-FEB-1997, integrated into UniProtKB/TrEMBL.
DT 01-FEB-1997, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Lys-gingipain.
GN Name=kgp;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
RA Prochazka V., Kiefer M.C., Travis J., Barr P.J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U54691; AAA99810.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 2.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
SQ SEQUENCE 1723 AA; 186832 MW; 4508A7E50197CEBD CRC64;

Query Match 94.2%; Score 114; DB 2; Length 1723;
Best Local Similarity 90.9%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||||| |||||
Db 432 LNTGVGFANYTAHGETSWADP 453

RESULT 8
CPG2_PORGI
ID CPG2_PORGI STANDARD; PRT; 736 AA.
```

```
DR EMBL; D83258; BAA11870.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 3.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
SQ SEQUENCE 1723 AA; 187262 MW; 5628963D251493EB CRC64;

Query Match 94.2%; Score 114; DB 2; Length 1723;
Best Local Similarity 90.9%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||||| |||||
Db 432 LNTGVGFANYTAHGETSWADP 453

RESULT 7
P72197_PORGI
ID P72197_PORGI PRELIMINARY; PRT; 1723 AA.
AC P72197;
DT 01-FEB-1997, integrated into UniProtKB/TrEMBL.
DT 01-FEB-1997, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Lys-gingipain.
GN Name=kgp;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
RA Prochazka V., Kiefer M.C., Travis J., Barr P.J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U54691; AAA99810.1; -; Genomic_DNA.
DR MEROPS; C25.002; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR Pfam; PF07675; Cleaved Adhesin; 2.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
SQ SEQUENCE 1723 AA; 186832 MW; 4508A7E50197CEBD CRC64;

Query Match 94.2%; Score 114; DB 2; Length 1723;
Best Local Similarity 90.9%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LNTGVSFANYTAHGETAWADP 22
||||| ||||||| |||||
Db 432 LNTGVGFANYTAHGETSWADP 453

RESULT 8
CPG2_PORGI
ID CPG2_PORGI STANDARD; PRT; 736 AA.
```


AC P95493; O33441;
 DT 15-DEC-1998, integrated into UniProtKB/Swiss-Prot.
 DT 03-OCT-2003, sequence version 2.
 DT 07-MAR-2006, entry version 53.
 DE Gingipain R2 precursor (EC 3.4.22.37) (Gingipain 2) (Arg-gingipain)
 DE (RGP-2).
 DE Name: rgpB; Synonyms: ptrlII, rgp2; OrderedLocusNames=PG0506;
 GN Porphyromonas gingivalis (Bacteroides gingivalis).
 OS Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RN NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PROTEIN SEQUENCE OF 230-651.
 RC STRAIN=HG66;
 RX MEDLINE=98370998; PubMed=9705298; DOI=10.1074/jbc.273.34.21648;
 RA Potempa J., Mikolajczyk-Pawlinska J., Brassell D., Nelson D.,
 RA Thøgersen I.B., Enghild J.J., Travis J.;
 RT "Comparative properties of two cysteine proteinases (gingipains R),
 RT the products of two related but individual genes of Porphyromonas
 RT gingivalis.";
 RL J. Biol. Chem. 273:21648-21657(1998).
 RN [2]
 RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=W50;
 RX MEDLINE=98304082; PubMed=9639929;
 RA Slakeski N., Bhogal P.S., O'Brien-Simpson N.M., Reynolds E.C.;
 RT "Characterization of a second cell-associated Arg-specific cysteine
 RT proteinase of Porphyromonas gingivalis and identification of an
 RT adhesin-binding motif involved in association of the prtr and prtK
 RT proteinases and adhesins into large complexes.";
 RL Microbiology 144:1583-1592(1998).
 RN [3]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=W83;
 RX MEDLINE=22829867; PubMed=12949112;
 RX DOI=10.1128/JB.185.18.5591-5601.2003;
 RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
 RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
 RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
 RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
 RA Dewhirst F.E., Fraser C.M.;
 RT "Complete genome sequence of the oral pathogenic bacterium
 RT Porphyromonas gingivalis strain W83.";
 RL J. Bacteriol. 185:5591-5601(2003).
 RN [4]
 RN ENZYME REGULATION.
 RX PubMed=11179305; DOI=10.1128/IAI.69.3.1402-1408.2001;
 RA Gusman H., Travis J., Helmerhorst E.J., Potempa J., Troxler R.F.,
 RA Oppenheim F.G.;
 RT "Salivary histatin 5 is an inhibitor of both host and bacterial
 RT enzymes implicated in periodontal disease.";
 RL Infect. Immun. 69:1402-1408(2001).
 RN [5]
 RN X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
 RX MEDLINE=99452753; PubMed=10523290; DOI=10.1093/emboj/18.20.5453;
 RA Eichinger A., Beisel H.-G., Jacob U., Huber R., Medrano F.-J.,
 RA Bandula A., Potempa J., Travis J., Bode W.;
 RT "Crystal structure of gingipain R: an Arg-specific bacterial cysteine
 RT proteinase with a caspase-like fold.";
 RL EMBO J. 18:5453-5462(1999).
 CC -!- FUNCTION: Thiol protease which is believed to participate in
 CC intracellular degradation and turnover of proteins. Its
 CC proteolytic activity is a major factor in both periodontal tissue
 CC destruction and in bacterial host defense mechanisms. Activates
 CC complement C3 and C5 (By similarity).
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins and small molecule
 CC substrates, with a preference for Arg in P1.
 CC -!- ENZYME REGULATION: Inhibited by human histatin-3 1/24 (histatin-
 CC 5).
 CC -!- SIMILARITY: Belongs to the peptidase C25 family.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC ----- U85038; AAB41892.1; -, Genomic DNA.
 DR EMBL; AF007124; AAC26371.1; -, Genomic DNA.
 DR EMBL; AE015924; AAO65700.1; -, Genomic DNA.
 DR PDB; 1CVR; X-ray; A=230-664.
 DR MEROPS; C25.003; -.
 DR GenomeReviews; AE015924_GR; PG0506.
 DR TIGR; PG0506; -.
 DR BioCyc; PGIN242619; PG0506-MONOMER; -.
 DR InterPro; IPR001769; Peptidase_C25.
 DR InterPro; IPR005536; Peptidase_C25_C.
 DR InterPro; IPR012600; Propeptide_C25.
 DR Pfam; PF01364; Peptidase_C25; 1.
 DR Pfam; PF03785; Peptidase_C25; 1.
 DR Pfam; PF08126; Propeptide_C25; 1.
 DR 3D-structure; Calcium; Complete proteome; Direct protein sequencing;
 KW Hydrolase; Protease; Signal; Thiol protease; Virulence; Zymogen.
 FT SIGNAL 1 24
 FT PROPEP 25 229
 FT FT
 FT CHAIN 230 736 /FTid=PRO_0000026535.
 FT Gingipain R2.
 FT //FTid=PRO_0000026536.
 FT Proton donor.
 FT ACT_SITE 440 Nucleophile.
 FT ACT_SITE 473 G -> D (in Ref. 1).
 FT CONFLICT 58 P -> A (in Ref. 1).
 FT CONFLICT 246 E -> G (in Ref. 1).
 FT CONFLICT 251 E -> K (in Ref. 1).
 FT CONFLICT 254 I -> V (in Ref. 1).
 FT CONFLICT 398 A -> V (in Ref. 1).
 FT CONFLICT 435 YNV -> FSM (in Ref. 1).
 FT CONFLICT 480 N -> D (in Ref. 1).
 FT CONFLICT 510 S -> Y (in Ref. 1).
 FT CONFLICT 512 S -> P (in Ref. 1).
 FT CONFLICT 515 K -> N (in Ref. 1).
 FT CONFLICT 560 K -> E (in Ref. 1).
 FT CONFLICT 582
 FT TURN 237 238
 FT STRAND 240 245
 FT HELIX 247 252
 FT HELIX 254 262
 FT TURN 263 264
 FT STRAND 266 271
 FT HELIX 272 275
 FT STRAND 277 278
 FT HELIX 281 293
 FT STRAND 294 294
 FT TURN 295 296
 FT STRAND 299 306
 FT TURN 308 310
 FT STRAND 311 311
 FT STRAND 314 316
 FT TURN 317 318
 FT STRAND 319 322
 FT HELIX 323 326
 FT STRAND 329 333
 FT STRAND 336 342
 FT STRAND 345 346
 FT HELIX 347 362
 FT STRAND 363 363
 FT TURN 366 367
 FT TURN 369 372
 FT STRAND 373 378
 FT STRAND 380 380
 FT STRAND 382 382
 FT TURN 384 385
 FT STRAND 386 386
 FT HELIX 387 389
 FT STRAND 391 391
 FT HELIX 392 406
 FT TURN 407 407
 FT STRAND 409 418
 FT HELIX 421 430
 FT STRAND 433 439

```
FT STRAND 441 441
FT TURN 443 446
FT TURN 447 449
FT STRAND 452 452
FT TURN 453 455
FT HELIX 456 458
FT TURN 462 463
FT STRAND 467 474
FT TURN 475 476
FT TURN 478 479
FT STRAND 480 482
FT HELIX 485 491
FT STRAND 493 494
FT TURN 495 496
FT STRAND 497 498
FT STRAND 500 500
FT STRAND 502 509
FT STRAND 513 514
FT HELIX 515 527
FT TURN 528 529
FT STRAND 531 531
FT TURN 532 533
FT STRAND 538 538
FT HELIX 539 564
FT STRAND 565 569
FT TURN 571 572
FT STRAND 574 574
FT STRAND 578 578
FT STRAND 581 581
FT STRAND 585 586
FT STRAND 589 592
FT TURN 593 594
FT STRAND 596 604
FT TURN 605 606
FT STRAND 608 613
FT TURN 614 615
FT STRAND 616 622
FT STRAND 624 624
FT TURN 625 626
FT STRAND 627 631
FT STRAND 633 634
FT TURN 637 638
FT STRAND 640 647
FT TURN 649 650
FT STRAND 651 651
FT STRAND 654 661
SQ SEQUENCE 736 AA; 80967 MW; C848DD3FAB420833 CRC64;

Query Match 67.8%; Score 82; DB 1; Length 736;
Best Local Similarity 77.8%; Pred. No. 0.00069;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NTGVSFANYTAHGSETAW 19
Db 429 NGGISLANYTGHGSETAW 446
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |

RESULT 9
Q51838_PORGI PRELIMINARY; PRT; 1706 AA.
ID Q51838_PORGI PRELIMINARY; PRT; 1706 AA.
AC Q51838;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1997, sequence version 2.
DT 07-FEB-2006, entry version 26.
DE Name=prpR1;
GN Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W50;

Query Match 67.8%; Score 82; DB 2; Length 1706;
Best Local Similarity 77.8%; Pred. No. 0.0017;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NTGVSFANYTAHGSETAW 19
Db 427 NGGISLANYTGHGSETAW 444
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |

RESULT 10
Q51839_PORGI PRELIMINARY; PRT; 1706 AA.
ID Q51839_PORGI PRELIMINARY; PRT; 1706 AA.
AC Q51839; Q51840;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE Arginine-specific thiol protease precursor.
GN Name=prtr;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W50;
RX MEDLINE=95160709; PubMed=7857299;
RA Kirszbaum L., Sotiropoulos C., Jackson C., Cleal S., Slakeski N.,
RA Reynolds E.C.;
RT "Complete nucleotide sequence of a gene prtr of Porphyromonas
RT gingivalis W50 encoding a 132 kDa protein that contains an arginine-
RT specific thiol endopeptidase domain and a haemagglutinin domain.";
RL Biochem. Biophys. Res. Commun. 207:424-431(1995).
RN [2]
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W50;
RX MEDLINE=96311139; PubMed=8713096; DOI=10.1006/bbrc.1996.1073;
RA Slakeski N., Cleal S.M., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene prrA that encodes
RL an arginine-specific thiol proteinase and multiple adhesins.";
RN Blochem. Biophys. Res. Commun. 224:605-610(1996).
CC [3]
RP NUCLEOTIDE SEQUENCE OF 212-455.
RC STRAIN=W50;
RA Reynolds E.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; L26341; AAC18876.1; -; Genomic_DNA.
DR HSSP; P95493; 1CVR.
DR SMR; Q51844; 228-655.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved_adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR InterPro; IPR012600; Propeptide_C25.
DR Pfam; PF07675; Cleaved_Adhesin; 2.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR Pfam; PF08126; Propeptide_C25; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW protease; Signal.
FT SIGNAL 1 23 Potential.
FT CHAIN 228 1706 arginine-specific thiol protease.
SQ SEQUENCE 1706 AA; 185627 MW; E8BDF07C9813B844 CRC64;

Query Match 67.8%; Score 82; DB 2; Length 1706;
Best Local Similarity 77.8%; Pred. No. 0.0017;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NTGVSFANYTAHGSETAW 19
|:|:|||||
Db 427 NGGISLANYTGHGSETAW 444

RESULT 11
ID Q7MTE2_PORGI PRELIMINARY; PRT; 1706 AA.
AC Q7MTE2
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 15-DEC-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Hemagglutinin protein HagE.
GN Name=hagE; OrderedLocusNames=PG2024; ORFNames=PG 2024;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
CC -----
```

```
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AE015924; AAC066991.1; -; Genomic_DNA.
DR HSSP; P95493; 1CVR.
DR SMR; Q7MTE2; 228-655.
DR TIGR; PG2024; -.
DR BiOCYC; PGIN242619; PG2024-MONOMER; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved_adhesin.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR InterPro; IPR012600; Propeptide_C25.
DR Pfam; PF07675; Cleaved_Adhesin; 2.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR Pfam; PF08126; Propeptide_C25; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1706 AA; 185673 MW; 6FE9B83AA98A2815 CRC64;

Query Match 67.8%; Score 82; DB 2; Length 1706;
Best Local Similarity 77.8%; Pred. No. 0.0017;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NTGVSFANYTAHGSETAW 19
|:|:|||||
Db 427 NGGISLANYTGHGSETAW 444

RESULT 12
Q51844_PORGI
ID Q51844_PORGI PRELIMINARY; PRT; 736 AA.
AC Q51844;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Arginine-specific cysteine proteinase (Arg-gingipain).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC33277;
RX MEDLINE=96007508; PubMed=7559528; DOI=10.1074/jbc.270.40.23619;
RA Nakayama K., Kadowaki T., Okamoto K., Yamamoto K.;
RT "Construction and characterization of arginine-specific cysteine
RT proteinase (Arg-gingipain)-deficient mutants of Porphyromonas
RT gingivalis. Evidence for significant contribution of Arg-gingipain to
RT virulence.";
RL J. Biol. Chem. 270:23619-23626(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC33277;
RX MEDLINE=97276476; PubMed=9130229;
RA Nakayama K.;
RT "Domain-specific rearrangement between the two Arg-gingipain-encoding
RT genes in Porphyromonas gingivalis: possible involvement of
RT nonreciprocal recombination.";
RL Microbiol. Immunol. 41:185-196(1997).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; D64081; BA010963.1; -; Genomic_DNA.
DR HSSP; P95493; 1CVR.
DR SMR; Q51844; 230-661.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
```

DR	InterPro; IPRO12600; Propeptide_C25.
DR	Pfam; PF01364; Peptidase_C25; 1.
DR	Pfam; PF03785; Peptidase_C25 C; 1.
DR	Pfam; PF08126; Propeptide_C25; 1.
KW	Calcium; Direct protein sequencing; Hydrolase; Protease; Signal;
KW	Thiol protease; Virulence; Zymogen.
FT	SIGNAL_1 24
FT	PROPEP 25 227
FT	CHAIN 228 991 /FTID=PRO_0000026533. Gingipain R1. /FTID=PRO_0000026534. Proton donor (By similarity). Nucleophile (By similarity). RT -> TK (in Ref. 2).
FT	ACT_SITE 438 438
FT	ACT_SITE 471 471
FT	CONFLICT 264 265
FT	SEQUENCE 991 AA; 108782 MW; 03EEF43CEBE2544 CRC64;
Query Match	64.5%; Score 78; DB 1; Length 991;
Best Local Similarity	72.2%; Pred. No. 0.0041; Mismatches 4; Indels 0; Gaps 0
MATCHES	13; Conservative 1; Mismatches 4; Indels 0; Gaps 0
QY	2 NTGVSFANYTAHGSSETAW 19
DB	427 NGGISLVNYTGHGSSETAW 444
RESULT 14	
ID	Q9R9B7_PORGI PRELIMINARY; PRT; 1687 AA.
AC	Q9R9B7;
DT	01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT	01-MAY-2000, sequence version 1.
DT	07-FEB-2006, entry version 19.
DE	Hemagglutinin/protease.
GN	Name=hage;
OS	Porphyromonas gingivalis (Bacteroides gingivalis).
OC	Bacteria; Bacteroidetes; Bacteroidales;
OC	Porphyromonadaceae; Porphyromonas.
NCBI_TaxID=837;	
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=381;
RA	Han N., Dong H., Progulske-Fox A.;
RL	Submitted (SRP-1997) to the EMBL/GenBank/DDBJ databases.
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms Distributed under the Creative Commons Attribution-NoDerivs License
EMBL; AF26946; AAD01810.1; -; Genomic_DNA.	
HSSP; P95493; 1CVR.	
SMR; Q9R9B7; 209-636.	
GO; GO:0008234; F:cysteine-type peptidase activity; IEA.	
GO; GO:0009405; P:pathogenesis; IEA.	
GO; GO:0006508; P:proteolysis; IEA.	
InterPro; IPR011628; Cleaved adhesin.	
InterPro; IPR000977; DNA ligase.	
InterPro; IPR01769; Peptidase_C25.	
InterPro; IPR005536; Peptidase_C25 C.	
InterPro; IPR012600; Peptidase_C25.	
Pfam; PF07675; Cleaved Adhesin; 2.	
Pfam; PF01364; Peptidase_C25; 1.	
Pfam; PF03785; Peptidase_C25 C; 1.	
Pfam; PF08126; Propeptide_C25; 1.	
PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.	
KW	Protease.
SEQUENCE	1687 AA; 193703 MW; D085B516A399FE70 CRC64;
Query Match	64.5%; Score 78; DB 2; Length 1687;
Best Local Similarity	72.2%; Pred. No. 0.0074; Mismatches 4; Indels 0; Gaps 0
MATCHES	13; Conservative 1; Mismatches 4; Indels 0; Gaps 0
QY	2 NTGVSFANYTAHGSSETAW 19
DB	408 NGGISLVNYTGHGSSETAW 425

Mon Aug 28 11:30:42 2006

```

RESULT 15
Q51816 PORGI
ID Q51816 PORGI PRELIMINARY; PRT; 1704 AA.
AC Q51816
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Arg-gingipain-1 proteinase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_taxid=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95138080; PubMed=7836351; DOI=10.1074/jbc.270.3.1007;
RA Pavloff N., Potempa J., Pike R.N., Prochazka V., Kiefer M.C.,
RA Travis J., Barr P.J.;
RT "Molecular cloning and structural characterization of the Arg-
RT gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
RT proteinase-adhesin polypeptide."
RL J. Biol. Chem. 270:1007-1010(1995).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U15282; AAA69539.1; -; Genomic_DNA.
DR PIR; A55426; A55426.
DR HSP; P95493; LCVR.
DR SMR; Q51816; 228-655.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR011628; Cleaved adhesin.
DR InterPro; IPR000977; DNA ligase.
DR InterPro; IPR001769; Peptidase_C25.
DR InterPro; IPR005536; Peptidase_C25_C.
DR InterPro; IPR012600; Propeptide_C25.
DR Pfam; PF07675; Cleaved Adhesin; 2.
DR Pfam; PF01364; Peptidase_C25; 1.
DR Pfam; PF03785; Peptidase_C25_C; 1.
DR Pfam; PF08126; Propeptide_C25; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
FT CHAIN 228 719 mature 50-kDa cysteine proteinase
FT CHAIN 719 gingipain.
SQ SEQUENCE 1704 AA; 185437 MW; 6A34B40131C2A676 CRC64;
Query Match 64.5%; Score 78; DB 2; Length 1704;
Best Local Similarity 72.2%; Pred. No. 0.0075;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 2 NTGVSFANYTAHGSETAW 19
Db 427 NGGISLVNYTGHGSETAW 444
```

Search completed: August 25, 2006, 18:04:21
Job time : 203.488 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:03:02 ; Search time 137 Seconds
(without alignments)
410.514 Million cell updates/sec

Title: US-10-018-892-3
Perfect score: 662
Sequence: 1 AAQNTTSANWSDPGFTGPA.....KSTLPAGTFTATFVQYQVN 128

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 21:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*
- 9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	662	100.0	128	AAB31431	Aab31431 Amino aci
2	662	100.0	144	AAB47082	Aab47082 Salmonell
3	662	100.0	165	AAW23571	Aaw23571 Salmonell
4	662	100.0	165	AAB31430	Aab31430 Amino aci
5	662	100.0	176	AAR23731	Aar23731 Fimbrial
6	662	100.0	176	AAR2173	Aar2173 SE fimbri
7	662	100.0	180	AAW47081	Aaw47081 Salmonell
8	653	98.6	165	AAW62752	Aaw62752 Seta sequ
9	93.5	14.1	431	3 AAG32324	Aag32324 Arabidops
10	93.5	14.1	431	3 AAG09613	Aag09613 Arabidops
11	93.5	14.1	431	8 ADI43535	Adi43535 Plant tra
12	93.5	14.1	431	8 AEA26765	Aea26765 Stress to
13	93.5	14.1	435	3 AAG09612	Aag09612 Arabidops
14	93.5	14.1	440	3 AAG49463	Aag49463 Arabidops
15	93.5	14.1	443	3 AAG49462	Aag49462 Arabidops
16	91.5	13.8	2468	6 ABP59933	Abp59933 Microbial
17	91.5	13.8	2468	6 ABP59933	Abp59933 Microbial
18	88.5	13.4	439	8 ADY25088	Ady25088 Plant ful
19	87.5	13.2	328	7 AAR39298	Aar39298 Maize dwa
20	86.5	13.1	2736	7 ABO81481	Abos81481 Pseudomon
21	83.5	12.6	219	8 ADY22824	Ady22824 Plant ful
22	83.5	12.6	346	8 ADJ34954	Adj34954 Xylanase
23	83	12.5	524	6 ABU34449	Abu34449 Protein e
24	83	12.5	549	4 AAG81139	Aag81139 Mycobacte

25	83	12.5	549	5	ABJ04694	Abj04694 Mycobacte
26	83	12.5	549	6	ABU36587	Abu36587 Protein e
27	83	12.5	660	2	AAV01303	Aav01303 Human tro
28	82.5	12.5	1415	7	ABO77692	Abos77692 Pseudomon
29	82.5	12.5	2763	2	AAW10344	Aaw10344 Maize dwa
30	82	12.4	347	8	ADJ35006	Adj35006 Xylanase
31	81.5	12.3	693	7	ABO70421	Abos70421 Pseudomon
32	81.5	12.3	1373	5	ABBS7364	Abbs7364 Mouse isc
33	81.5	12.3	1373	9	ADM44460	Adm44460 Murine pr
34	81.5	12.3	1721	6	ABU34624	Abu34624 Protein e
35	81.5	12.3	2204	6	ABU36440	Abu36440 Protein e
36	81.5	12.3	2204	9	ABBS1424	Abbs1424 Microbial
37	81	12.2	329	8	ABOS8582	Abos8582 Human gen
38	81	12.2	334	6	ABU20015	Abu20015 Protein e
39	81	12.2	605	6	ADA55641	Ada55641 Human pro
40	81	12.2	827	6	ABU19388	Abu19388 Protein e
41	80.5	12.2	558	8	ADU07619	Adu07619 Amino aci
42	80.5	12.2	595	6	ABU19799	Abu19799 Protein e
43	80.5	12.2	999	7	ADJ70227	Adj70227 Human hea
44	80.5	12.2	1862	6	ABU21833	Abu21833 Protein e
45	79.5	12.0	570	8	ADQ65871	Adq65871 Novel hum

ALIGNMENTS

RESULT 1
AAB31431
ID AAB31431 standard; peptide; 128 AA.
AC AAB31431;
DT 20-APR-2001 (first entry)
DE Amino acid sequence of the C128 fragment of the Sef14 antigen.
KW C128 fragment; fimbrial Sef14 antigen; fimbrial protein;
KW flagellin protein; poultry.
OS Salmonella enteritidis.

XX WO200078995-A1.
XX 28-DEC-2000.
XX 22-JUN-1999; 99WO-SG000061.
XX 22-JUN-1999; 99WO-SG000061.
(MOLE-) INST MOLECULAR AGROBIOLOGY.
XX Kwang H, Liu W, Low SS, Loh KYH;
XX WPI; 2001-071400/08.
PT New method for the specific detection of Salmonella enteritidis
PT infections of poultry comprises contacting a biological sample with
PT antigenic fragments of S. enteritidis fimbrial and/or flagellin proteins.
Claim 18; Page 42; 49pp; English.

XX The present sequence represents the C128 fragment of the fimbrial Sef14
XX antigen of Salmonella enteritidis. The specification describes a method
XX for detecting S. enteritidis in a biological sample obtained from
XX poultry. The method comprises contacting the sample with an antigenic
XX fragment of S. enteritidis fimbrial or flagellin protein and detecting
XX the formation of a complex, where the fragment is specifically recognized
XX by S. enteritidis antibodies. The antigenic fragments are specific to
XX Salmonella enteritidis and enable specific detection of S. enteritidis
XX even in the presence of other Salmonella spp. The methods are useful for
XX the specific detection of S. enteritidis infections in biological samples
XX derived from poultry

SQ Sequence 128 AA;
 Query Match 100.0%; Score 662; DB 4; Length 128;
 Best Local Similarity 100.0%; Pred. No. 3e-59;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAOQTTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
 DB 1 AAOQTTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60

QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
 DB 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120

QY 121 FYVQOYQN 128
 DB 121 FYVQOYQN 128

RESULT 2
 AAW47082
 ID AAW47082 standard; protein; 144 AA.
 AC AAW47082;
 XX
 DT 25-JUN-1998 (first entry)
 DE
 KW Salmonella Sef14 protein fragment.
 XX
 KW Sefa gene; Sef14 protein; infection; bird; chicken; turkey;
 KW anti-Salmonella enteritidis antibody; vaccine; poultry.
 XX
 OS Salmonella enteritidis.
 XX
 FN W09803656-A1.
 XX
 PD 29-JAN-1998.
 PF 18-JUL-1997; 97MO-US012639.
 XX
 PR 19-JUL-1996; 96US-0022191P.
 XX
 PA (MINU) UNIV MINNESOTA.
 XX
 PI Rajashekara G, Nagaraja KV, Kapur V;
 XX
 DR WPI; 1998-120780/11.
 N-PSDB; AAV13948.
 XX
 PT Detecting antibodies against Salmonella enteritidis using truncated
 PT fimbrial antigen Sef14 - in immunoassays, particularly for diagnosing
 PT infection in poultry, also new antigens.
 XX
 PS Claim 4; Page 23-24; 38pp; English.
 XX
 CC This sequence is a fragment of the Salmonella enteritidis (Se) Sef14
 CC protein, encoded by the sefa gene. The method of the invention is for
 CC detecting antibodies (Ab) against Se in an animal by treating a sample
 CC with a truncated Sef14 antigen (Ag), lacking at least the native Sef14
 CC signal peptide, and detecting any Ab-Ag complex formed. Detection (by
 CC enzyme-linked immunosorbent assay or agglutination tests) of the Ab is
 CC used to diagnose Se infection in birds, especially chickens and turkeys.
 CC The Ag can also be used in vaccines to protect poultry against Se
 CC infection. Detection of the Ab is a sensitive, specific method for
 CC reliable and routine screening of animals. The Ag are easily produced in
 CC large quantities, in pure form, without requiring special growing
 CC conditions, so are suitable for large scale screening of flocks
 XX
 SQ Sequence 144 AA;
 Query Match 100.0%; Score 662; DB 2; Length 144;
 Best Local Similarity 100.0%; Pred. No. 3.5e-59;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAOQTTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
 DB 17 AAOQTTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 76

QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
 DB 77 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 136

QY 121 FYVQOYQN 128
 DB 137 FYVQOYQN 144

RESULT 3
 AAW23571
 ID AAW23571 standard; protein; 165 AA.
 AC AAW23571;
 XX
 DT 25-MAR-2003 (revised)
 DT 29-SEP-1997 (first entry)
 XX
 DE Salmonella enteritidis sefa.
 XX
 KW Enteropathogenic bacteria; enterobacteria; S. enteritidis; antibody.
 XX
 OS Salmonella enteritidis.
 XX
 FN US5635617-A.
 PD 03-JUN-1997.
 XX
 PF 26-APR-1994; 94US-00233788.
 XX
 PR 26-APR-1993; 93US-00054452.
 XX
 PA (UYVI-) UNIV VICTORIA INNOVATION & DEV CORP.
 XX
 PI Collinson SK, Kay WM, Doran JL;
 XX
 DR WPI; 1997-309886/28.
 N-PSDB; AAT74143.
 XX
 PT Isolated Salmonella gene agfa - used for diagnosis of Salmonella or
 PT enteropathogenic bacteria of the Enterobacteria family.
 XX
 PS Example 9; Fig 2; 85pp; English.
 XX
 CC The present sequence represents sefa found in the sef gene cluster from
 CC Salmonella enteritidis. The nucleic acid can be used to provide
 CC diagnostic assays for Salmonella and/or enteropathogenic bacteria of the
 CC family Enterobacteria. It can also be used to provide proteins and
 CC antibodies which can be used for assays. The nucleic acid sequence can be
 CC used to provide probes or primers which can specifically hybridize to
 CC nucleic acid molecules from greater than 99% of Salmonella strains that
 CC are pathogenic to warm-blooded animals relative to nucleic acid molecules
 CC from virtually all other microbial organisms. (Updated on 25-MAR-2003 to
 CC correct PF field.)
 XX
 SQ Sequence 165 AA;
 Query Match 100.0%; Score 662; DB 2; Length 165;
 Best Local Similarity 100.0%; Pred. No. 4.1e-59;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAOQTTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
 DB 38 AAOQTTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 97

QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
 DB 98 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 157

Qy 121 FYVQYQN 128
 |||||
 Db 158 FYVQYQN 165

RESULT 4
 AAB31430
 ID AAB31430 standard; protein; 165 AA.

XX AAB31430;

XX 20-APR-2001 (first entry)

XX Amino acid sequence of the Sef14 antigen.

XX C128 fragment; fimbrial Sef14 antigen; fimbrial protein;
 KW flagellin protein; poultry.

XX Salmonella enteritidis.

OS WO200078995-A1.

PN 28-DEC-2000.

XX 22-JUN-1999; 99WO-SG000061.

XX 22-JUN-1999; 99WO-SG000061.

XX (MOLE-) INST MOLECULAR AGROBIOLOGY.

PA Kwang H, Liu W, Low SS, Loh KYH;

PI WPI; 2001-071400/08.

XX N-PSDB; AAF24784.

XX New method for the specific detection of Salmonella enteritidis
 PT infections of poultry comprises contacting a biological sample with
 PT antigenic fragments of S. enteritidis fimbrial and/or flagellin proteins.

XX Disclosure; Page 42; 49pp; English.

XX The present sequence represents the fimbrial Sef14 antigen of Salmonella
 CC enteritidis. The specification describes a method for detecting S.
 CC enteritidis in a biological sample obtained from poultry. The method
 CC comprises contacting the sample with an antigenic fragment of S.
 CC enteritidis fimbrial or flagellin protein and detecting the formation of
 CC a complex, where the fragment is specifically recognized by S.
 CC enteritidis antibodies. The antigenic fragments are specific to
 CC Salmonella enteritidis and enable specific detection of S. enteritidis
 CC even in the presence of other Salmonella spp. The methods are useful for
 CC the specific detection of S. enteritidis infections in biological samples
 CC derived from poultry

SQ Sequence 165 AA;

Query Match 100.0%; Score 662; DB 4; Length 165;

Best Local Similarity 100.0%; Pred. No. 4.1e-59;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAOQNTTSANWSQDPGFTGPAVAAGQKVGTLTATGPHNSVSIAGKASVSGGVATVPFV 60
 |||||
 Db 38 AAOQNTTSANWSQDPGFTGPAVAAGQKVGTLTATGPHNSVSIAGKASVSGGVATVPFV 97
 |||||

Qy 61 DGCGQPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120

Db 98 DGCGQPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 157

Qy 121 FYVQYQN 128

Db 158 FYVQYQN 165

RESULT 5

AAR23731

ID AAR23731 standard; protein; 176 AA.

XX AAR23731;

XX 25-MAR-2003 (revised)

DT 02-NOV-1992 (first entry)

XX Fimbrial antigen.

XX S. dublin; S. moscow; fimbria-like structure; epitope.

XX Salmonella enteritidis.

PN WO9206198-A.

XX 16-APR-1992.

XX 01-OCT-1991; 91WO-GB001691.

PR 01-OCT-1990; 90GB-00021338.

PR 17-OCT-1990; 90GB-00022570.

XX (UKAG-) UK MIN AGRIC FISH.

XX Woodward MJ;

XX WPI; 1992-150883/18.

XX Detection and identification of salmonella - by using monoclonal
 PT antibodies to detect epitope(s) of these serotypes in culture.

XX Disclosure; Page 3; 48pp; English.

XX The sequence given is Salmonella enteritidis fimbrial antigen (SEFA).
 CC Salmonella organisms have fimbria-like structures on their surfaces and
 CC it has been suggested that there are antigenically distinct types of
 CC fimbria, ie. possessing specific epitopes on the fimbrial antigens. This
 CC sequence has an amino acid sequence which forms an epitope on the fimbria
 CC "in vivo" which is specifically encoded by DNA of the species S.
 CC enteritidis, and some strains of the species S. dublin and S. moscow but
 CC which is apparently absent in virtually all other serotypes. This
 CC antigen can be used for testing for the presence of Salmonella
 CC microorganisms in clinical samples such as animal remains or prods., food
 CC samples and infected environmental samples. (Updated on 25-MAR-2003 to
 CC correct PF field.)

SQ Sequence 176 AA;

Query Match 100.0%; Score 662; DB 2; Length 176;

Best Local Similarity 100.0%; Pred. No. 4.4e-59;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAOQNTTSANWSQDPGFTGPAVAAGQKVGTLTATGPHNSVSIAGKASVSGGVATVPFV 60
 |||||

Db 49 AAOQNTTSANWSQDPGFTGPAVAAGQKVGTLTATGPHNSVSIAGKASVSGGVATVPFV 108
 |||||

Qy 61 DGCGQPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120

Db 109 DGCGQPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 168

Qy 121 FYVQYQN 128

Db 169 FYVQYQN 176

RESULT 6

AAR42173

ID AAR42173 standard; protein; 176 AA.

XX AAR42173;

DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 03-MAY-1994 (first entry)
 XX
 XX SE fimbrial antigen (SEFA).
 DE Salmonella enteritidis fimbrial antigen; SEFA; Salmonella typhi;
 KW Salmonella dublin; serotype.
 XX
 OS Salmonella enteritidis.
 OS Salmonella typhi.
 XX WO9320231-A1.
 XX 14-OCT-1993.
 XX 29-MAR-1993; 93WO-GB000647.
 XX 31-MAR-1992; 92GB-00007069.
 XX (UKAG-) UK MIN FISHERIES & FOOD.
 XX Woodward MJ, Thorns CJ;
 XX WPI; 1993-336937/42.
 DR N-PSDB; AAQ49882.
 XX Testing for Salmonella serotypes, esp. S. Typhi - using test kit for
 PT detecting nucleic acid sequences specific to certain sero-types.
 XX Claim 2; Page 19-21; 37pp; English.
 XX DNA encoding SEFA is common to members of the enteritidis, dublin and
 CC typhi serogroups and can therefore be used in the detection of such
 CC organisms. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-
 CC AUG-2003 to correct OS field.)
 XX Sequence 176 AA;
 SQ
 Query Match 100.0%; Score 662; DB 2; Length 176;
 Best Local Similarity 100.0%; Pred. No. 4.4e-59;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLTITATGPHNSVSIAGKGSVSGGVATVPFV 60
 DB 49 AAQNTTSANWSQDPGFTGPAVAAGKVGTLTITATGPHNSVSIAGKGSVSGGVATVPFV 108
 QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
 DB 109 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 168
 QY 121 FYVQOYQN 128
 DB 169 FYVQOYQN 176
 RESULT 7
 AAW47081
 ID AAW47081 standard; protein; 180 AA.
 XX
 AC AAW47081;
 XX 25-JUN-1998 (first entry)
 DT Salmonella Sef14 protein fragment.
 DE Sefa gene; Sef14 protein; infection; bird; chicken; turkey;
 KW anti-Salmonella enteritidis antibody; vaccine; poultry.
 XX Salmonella enteritidis.
 OS WO9803656-A1.
 XX

PD 29-JAN-1998.
 XX 18-JUL-1997; 97WO-US012639.
 XX 19-JUL-1996; 96US-0022191P.
 XX (MINU) UNIV MINNESOTA.
 XX Rajashekara G, Nagaraja KV, Kapur V;
 FI WPI; 1998-120780/11.
 XX N-PSDB; AAV13974.
 XX Detecting antibodies against Salmonella enteritidis using truncated
 PT fimbrial antigen Sef14 - in immunoassays, particularly for diagnosing
 PT infection in poultry, also new antigens.
 XX Claim 3; Page 21-22; 38pp; English.
 XX This sequence is a fragment of the Salmonella enteritidis (Se) Sef14
 CC protein, encoded by the sefa gene. The method of the invention is for
 CC detecting antibodies (Ab) against Se in an animal by treating a sample
 CC with a truncated Sef14 antigen (Ag), lacking at least the native Sef14
 CC signal peptide, and detecting any Ab-Ag complex formed. Detection (by
 CC enzyme-linked immunosorbent assay or agglutination tests) of the Ab is
 CC used to diagnose Se infection in birds, especially chickens and turkeys.
 CC The Ag can also be used in vaccines to protect poultry against Se
 CC infection. Detection of the Ab is a sensitive, specific method for
 CC reliable and routine screening of animals. The Ag are easily produced in
 CC large quantities, in pure form, without requiring special growing
 CC conditions, so are suitable for large scale screening of flocks
 XX Sequence 180 AA;
 SQ
 Query Match 100.0%; Score 662; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 4.6e-59;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLTITATGPHNSVSIAGKGSVSGGVATVPFV 60
 DB 53 AAQNTTSANWSQDPGFTGPAVAAGKVGTLTITATGPHNSVSIAGKGSVSGGVATVPFV 112
 QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
 DB 113 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 172
 QY 121 FYVQOYQN 128
 DB 173 FYVQOYQN 180
 RESULT 8
 AAR62752
 ID AAR62752 standard; protein; 165 AA.
 XX
 AC AAR62752;
 XX 21-OCT-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 26-JUN-1995 (first entry)
 XX Sefa sequence.
 DE Salmonella; Sefa; vaccine.
 KW Salmonella sp.
 OS Unidentified.
 OS WO9425598-A2.
 XX 10-NOV-1994.
 PD 26-APR-1994; 94WO-IB000207.
 XX

```
XX PR 26-APR-1993; 93US-00054452.
XX PR (UVI-) UNIV VICTORIA INNOVATION & DEV CORP.
PA PA (KING/) KING J.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
PI PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
DR DR N-PSDB; AAQ73061.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 06-MAY-1999; 99US-0132487P.
PT PR 07-MAY-1999; 99US-0132863P.
PT PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.
CC PR 18-MAY-1999; 99US-0134768P.
CC PR 19-MAY-1999; 99US-0134941P.
CC PR 20-MAY-1999; 99US-0135124P.
CC PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0135629P.
XX PR 25-MAY-1999; 99US-0136021P.
SQ Sequence 165 AA;
Query Match 98.6%; Score 653; DB 2; Length 165;
Best Local Similarity 99.2%; Pred. No. 3.4e-58;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AAQNTTSANWSQDPGFTGPAVAAGKQVGTLSITATGPHNSVSIACKGASVSGGVATVPV 60
Db |||||
38 AAQNTTSANWSQDPGFTGPAVAAGKQVGTLSITATGPHNSVSIACKGASVSGGVATVPV 97
Qy 61 DGGQPVFRGRIQGANINDQANTGIDGLAGRWASSQETLNVPVTFCKSTLPAGTFTTAT 120
Db |||||
98 DFGQPVFRGRIQGANINDQANTGIDGLAGRWASSQETLNVPVTFCKSTLPAGTFTTAT 157
Qy 121 FYVQQYQN 128
Db |||||
158 FYVQQYQN 165
RESULT 9
AAG32324
ID AAG32324 standard; protein; 431 AA.
XX AC AAG32324;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 38974.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
19-APR-1999; 99US-0130077P.
21-APR-1999; 99US-0130449P.
23-APR-1999; 99US-0130510P.
23-APR-1999; 99US-0130891P.
28-APR-1999; 99US-0131449P.
30-APR-1999; 99US-0132048P.
30-APR-1999; 99US-0132407P.
04-MAY-1999; 99US-0132484P.
05-MAY-1999; 99US-0132485P.
06-MAY-1999; 99US-0132486P.
06-MAY-1999; 99US-0132487P.
07-MAY-1999; 99US-0132863P.
11-MAY-1999; 99US-0134256P.
14-MAY-1999; 99US-0134218P.
14-MAY-1999; 99US-0134221P.
14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
```

PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147431P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158316P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 25-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 14.1%; Score 93.5; DB 3; Length 431;
 Best Local Similarity 33.3%; Pred. No. 0.94; Indels 11; Gaps 6;
 Matches 36; Conservative 14; Mismatches 47

QY 1 AAQNTTSANWSQDPGTGPAVAAGQKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
 DB 68 ATANTTTASSSDSPSSAAAAAANQWLSSSSFLQNNNNNASIVGDGDDVTGGADTMI 127
 QY 59 FVDGQGPVFRGRIQGANINDQAN-TGIDGLAGRWASSQ-ETLVNVPV 104
 DB 128 ---QGENKMTGG---GENKNDGGGATAADGVVSWQNAHKAEILSHPL 168

RESULT 10
 AAG09613
 ID AAG09613 standard; protein; 431 AA.
 XX AAG09613;
 AC AAG09613;
 XX 17-OCT-2000 (first entry)
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 7614.
 DE Arabidopsis thaliana protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX EP1033405-A2.
 PN 06-SEP-2000.
 PD 25-FEB-2000; 2000EP-00301439.
 PF 25-FEB-1999; 99US-0121825P.
 XX 05-MAR-1999; 99US-0123180P.
 PR 03-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.

PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159884P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 14.1%; Score 93.5; DB 3; Length 431;
 Best Local Similarity 33.3%; Pred. No. 0.94;
 Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;
 QY 1 AAQNTTSANWSQDPGFTGPAVAGQKVG-TLSITATGPHNSVSIAGKA-SVSGGVATVP 58
 Db 68 ATANTTTASSSSDPSSAAAAAANQWLSSFLQRNNNNASIVGDGIDDDVTGGADTMI 127
 QY 59 FVDSGQGVFRGRIQGANINDQAN-TGIDGLAGRWVASSQ-ETLNVV 104
 Db 128 ----QGEKMTGG---GENKNDGGGATAADGVVSWQNAHKAEILSHPL 168

RESULT 11
 ADI43535
 ID ADI43535 standard; protein; 431 AA.
 XX
 AC ADI43535;
 XX
 DT 22-APR-2004 (first entry)
 XX

Plant transcription factor #640.
 DE transgenic; plant; enhanced tolerance to abiotic stress;
 KW glyphosate tolerance; hormone sensitivity; disease resistance;
 KW sugar sensing; flowering; flower structure; stem bifurcation;
 KW branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; gene; ds.

OS Arabidopsis thaliana.
 XX
 XX US2004019927-A1.
 PN
 XX 29-JAN-2004.
 XX
 XX 25-FEB-2003; 2003US-00374780.
 XX
 XX 18-APR-2001; 2001US-00837944.
 XX

(SHER/) SHERMAN B K.
 PA (RIEC/) RIECHMANN J L.
 PA (JIAN/) JIANG C.
 PA (HEAR/) HEARD J E.
 PA (HAKE/) HAAKE V.
 PA (CREE/) CREELMAN R A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L J.

PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P E.
 PA (FILG/) FILGRIM M L.
 PA (DUBE/) DUBELL A N.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
 PI Filgrim ML, Dubell AN, Pineda O, Yu G;
 XX
 DR WPI; 2004-132245/13.

PT New transgenic plant comprising a recombinant polynucleotide of any one
 FT of more than 500 nucleotide sequences, useful in bioinformatic search
 FT methods.

PS Claim 1; SEQ ID NO 1998; 435pp; English.

XX The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produce a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyphosate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure; change in stem bifurcations; altered branching
 CC pattern; reduced apical dominance; reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins; or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This is the amino acid sequence of a plant
 CC transcription factor, and an orthologue of Arabidopsis thaliana
 CC transcription factor isolated in the invention, that can be used in the
 CC creation of a transgenic plant with altered traits.

SQ Sequence 431 AA;

Query Match 14.1%; Score 93.5; DB 8; Length 431;
 Best Local Similarity 33.3%; Pred. No. 0.94;
 Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;

QY 1 AAQNTTSANWSQDPGFTGPAVAGQKVG-TLSITATGPHNSVSIAGKA-SVSGGVATVP 58
 Db 68 ATANTTTASSSSDPSSAAAAAANQWLSSFLQRNNNNASIVGDGIDDDVTGGADTMI 127
 QY 59 FVDSGQGVFRGRIQGANINDQAN-TGIDGLAGRWVASSQ-ETLNVV 104
 Db 128 ----QGEKMTGG---GENKNDGGGATAADGVVSWQNAHKAEILSHPL 168

RESULT 12
 AEA26765
 ID AEA26765 standard; protein; 431 AA.
 XX
 AC AEA26765;
 XX

DT 28-JUL-2005 (first entry)
 XX
 DE Stress tolerant plant-related transcription factor protein SeqID606.
 XX transcription factor; transgenic plant; agriculture; drought resistance;
 KW stress tolerance.
 XX Arabidopsis thaliana.
 OS
 XX WO2005047516-A2.
 PN

PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 21-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 22-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149175P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149266P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149923P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151063P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154773P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156598P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158389P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160788P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160800P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 14.1%; Score 93.5; DB 3; Length 435;

Best Local Similarity 33.3%; Pred. No. 0.95;

Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;

Qy 1 AAQNTTSANWSQDPCFTGPAVAAGOKVG-TLSITATGPHNSVSIAGKA-SVSGGVATVP 58

Db 72 ATANTTTASSDSSPSSAAAAAANQWLSSFLQNNNNNASIVGGDDVTGGADTMI 131

Qy 59 FVDGQGPVFRIGQANINDQAN-TGIDGLAGWRVASSQ-ETLNVV 104

Db 132 ----QEMKTGG--GENKNDGGGATAADGVVSWQNAHKAILSHPL 172

RESULT 14

AAG49463

ID AAG49463 standard; protein; 440 AA.

XX AAG49463;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 62578.

DE Arabidopsis thaliana.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999;	99US-0126785P.	PR 19-JUL-1999;	99US-0144325P.
PR 01-APR-1999;	99US-0127462P.	PR 19-JUL-1999;	99US-0144331P.
PR 06-APR-1999;	99US-0128234P.	PR 19-JUL-1999;	99US-0144332P.
PR 08-APR-1999;	99US-0128714P.	PR 19-JUL-1999;	99US-0144333P.
PR 16-APR-1999;	99US-0128845P.	PR 19-JUL-1999;	99US-0144334P.
PR 19-APR-1999;	99US-0130077P.	PR 19-JUL-1999;	99US-0144335P.
PR 21-APR-1999;	99US-0130449P.	PR 20-JUL-1999;	99US-0144352P.
PR 23-APR-1999;	99US-0130510P.	PR 20-JUL-1999;	99US-0144352P.
PR 23-APR-1999;	99US-0130891P.	PR 20-JUL-1999;	99US-0144884P.
PR 28-APR-1999;	99US-0131448P.	PR 21-JUL-1999;	99US-0144814P.
PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145086P.
PR 30-APR-1999;	99US-0132407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-0132484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-0132485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-0132486P.	PR 22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999;	99US-0132487P.	PR 22-JUL-1999;	99US-0145192P.
PR 07-MAY-1999;	99US-0132863P.	PR 23-JUL-1999;	99US-0145145P.
PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135623P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140931P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 02-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 06-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 08-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 09-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 12-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.


```

PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0158294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161820P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 14.1%; Score 93.5; DB 3; Length 440;
Best Local Similarity 33.3%; Pred. No. 0.97;
Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;

QY 1 AAQNTTSANWSQDFGTGPAVAGQKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
DB 68 ATANTTTASSDSFSSAAAAANQWLRSRSSFLORNNNNNASIVGDIIDVTGADTMI 127
QY 59 FVDGQGFVRGRIOGANINDQAN-TGIDGLAGHRVASSQ-ETLNVVP 104
DB 128 ----QGMETGKT---GENKNDGGGATAADGVVSWQNAHRKAEILSHPL 168

RESULT 15
AAG49462
ID AAG49462 standard; protein; 443 AA.
XX
XX AAG49462;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 62577.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.

```

PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144633P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145091P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 14.1%; Score 93.5; DB 3; Length 443;
Best Local Similarity 33.3%; Pred. No. 0.98;

Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;

QY 1 AAQNTTSANWSQDPGFTGPAVAGQKVG-TLSITATGPHNSVSIAGKA-SVSGGVATVP 58

Db 71 ATANTTTASSDSPSSAAAAAANQWLRSSEFLQRNNNNASIVGDGDDVTGGADTMI 130

QY 59 FVDGQGPVFRGRIQANINDQAN-TGIDGLAGRWASSQ-ETLNVFV 104

Db 131 ----QGEKMTGG---GENKNDGGGATAADGVVSWQARHKAELSHPL 171

Search completed: April 26, 2006, 17:06:21

Job time : 139 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:02:58 ; Search time 74 Seconds
(without alignments)
1220.373 Million cell updates/sec

Title: US-10-018-892-3

Perfect score: 652

Sequence: 1 AAQNTTSANWSQDGFRTGPA.....KSTLPAGTFTATFYVQYQN 128

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	662	100.0	165	1 FM_SALEN	P12061 salmonella
2	637	96.2	165	2 Q5PM43 SALPA	Q5pm43 salmonella
3	295	44.6	166	2 Q9X6U1 ECOLI	Q9x6u1 escherichia
4	270	40.8	166	2 Q47405 ECOLI	Q47405 escherichia
5	98	14.8	656	2 Q50Q12 ENTHI	Q50q12 entamoeba h
6	94.5	14.3	1929	2 Q54JT4 DICDI	Q54j14 dictyosteli
7	93.5	14.1	431	1 KNAT73 ARATH	P48000 arabidopsis
8	93.5	14.1	431	2 Q8LDC1 ARATH	Q8ldc1 arabidopsis
9	93	14.0	398	2 Q6VA08 TREPA	Q6va08 treponema p
10	93	14.0	400	2 Q6VAL3 TREPA	Q6val3 treponema p
11	93	14.0	401	2 Q6VA23 TREPA	Q6va23 treponema p
12	93	14.0	405	2 Q6VA22 TREPA	Q6va22 treponema p
13	91.5	13.8	2468	2 Q9I2M3 PSEAB	Q9i2m3 pseudomonas
14	90.5	13.7	1973	2 Q7NP45 GLOVI	C7np45 gloeobacter
15	90	13.6	476	2 Q4J4Q1 AZOVI	Q4j4q1 azotobacter
16	88	13.3	405	2 Q6VAL9 TREPA	Q6val9 treponema p
17	88	13.3	499	2 Q9EUA1 TREPA	Q9eua1 treponema p
18	87.5	13.2	380	1 POLG MDWV	P32652 maize dwarf
19	87	13.1	394	2 Q6VA59 TREPA	Q6va59 treponema p
20	87	13.1	609	2 Q82K19 STRAW	Q82k19 streptomyc
21	87	13.1	1778	2 Q7SZY1 BRARE	Q7szy1 brachydanio
22	86.5	13.1	399	2 Q6V920 TREPA	Q6v920 treponema p
23	86.5	13.1	1356	1 CO1A2 ONCMY	O934a0 oncorhynch
24	86	13.0	1472	2 Q90ZAO CHICK	Q90za0 gallus gall
25	86	13.0	2751	2 Q8UKU0 RALSO	Q8uku0 ralstonia s
26	85.5	12.9	1346	2 Q8UJ3 ONCKE	Q8uu3 oncorhynch
27	85	12.8	173	2 Q6PFR1 ACTAD	Q6pfr1 acinetobact
28	85	12.8	400	2 Q6VA53 TREPA	Q6va53 treponema p
29	85	12.8	496	2 Q84AM4 TREPA	Q84am4 treponema p
30	85	12.8	496	2 Q84AM5 TREPA	Q84am5 treponema p
31	85	12.8	496	2 Q84AM6 TREPA	Q84am6 treponema p

RESULT 1

ID	FM_SALEN	STANDARD	PRT	165 AA.
AC	P12061			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Fimbrial protein precursor.			
GN	Name=sefa; Synonyms=sef14;			
OS	Salmonella enteritidis.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Salmonella.			
OX	NCBI_TaxID=592;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=27655-3B;			
RX	MEDLINE=93239677; PubMed=8097515;			
RA	Clouthier S.C., Mueller K.-H., Doran J.L., Collinson S.K., Kay W.W.;			
RT	"Characterization of three fimbrial genes, sefABC, of Salmonella enteritidis."			
RL	J. Bacteriol. 175:2523-2533 (1993).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=91072589; PubMed=1701443;			
RA	Thorns C.J., Sojka M.G., Chasey D.C.;			
RT	"Detection of a novel fimbrial structure on the surface of Salmonella enteritidis by using a monoclonal antibody."			
RL	J. Clin. Microbiol. 28:2409-2414 (1990).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Ogunniyi A.D., Kotlarski I., Morona R., Manning P.A.;			
RL	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	PROTEIN SEQUENCE OF 22-85.			
RX	MEDLINE=87008384; PubMed=2875990;			
RA	Feutrier J., Kay W.W., Trust T.J.;			
RT	"Purification and characterization of fimbriae from Salmonella enteritidis."			
RL	J. Bacteriol. 168:221-227 (1986).			
CC	-!- FUNCTION: Structural subunit of the sef14 fimbriae (S. enteritidis filamentous fimbriae).			
CC	-!- SUBCELLULAR LOCATION: Fimbria.			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
CC	-----			
CC	EMBL; L11008; AAA27219.1; -; Genomic DNA.			
DR	EMBL; L03833; AAA71892.1; -; Unassigned DNA.			
DR	EMBL; X98516; CAA67141.1; -; Genomic DNA.			
DR	PIR; A40618; A40618.			
DR	PDB; 1LUO; Model; A-1-165.			

```

DR InterPro; IPR010498; SEF14 adhesin.
DR Pfam; PF06443; SEF14 adhesin; 1.
KW 3D-structure; Direct protein sequencing; Fimbria; Signal.
FT SIGNAL 1 21
FT CHAIN 22 165 Fimbrial protein.
FT CONFLICT 30 30 V -> E (in Ref. 2 and 3).
FT CONFLICT 84 85 GA -> QW (in Ref. 4).
SQ SEQUENCE 165 AA; 16477 MW; 5B33798A3F0F9091 CRC64;

Query Match 100.0%; Score 662; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.4e-50;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
Db 38 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 97
QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPVTTFGKSTLPAGTTAT 120
Db 98 DGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPVTTFGKSTLPAGTTAT 157
QY 121 FYVQYQYQN 128
Db 158 FYVQYQYQN 165

RESULT 2
QSPM43 SALPA
ID QSPM43 SALPA PRELIMINARY; PRT; 165 AA.
AC QSPM43; 2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Fimbrial structural protein.
GN Name-sefa; OrderedLocName=SPM4304;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
DR NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P., Florea L.,
RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV80033.1; -; Genomic_DNA.
DR InterPro; IPR010498; SEF14 adhesin.
DR Pfam; PF06443; SEF14 adhesin; 1.
KW Complete proteome.
SQ SEQUENCE 165 AA; 16665 MW; 8A32BE3F43C91520 CRC64;

Query Match 96.2%; Score 637; DB 2; Length 165;
Best Local Similarity 96.1%; Pred. No. 8.3e-48;
Matches 123; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
Db 38 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 97
QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPVTTFGKSTLPAGTTAT 120
Db 98 DGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPVTTFGKSTLPAGTTAT 157
QY 121 FYVQYQYQN 128
Db 158 FYVQYQYQN 165

Query Match 44.6%; Score 295; DB 2; Length 166;
Best Local Similarity 45.8%; Pred. No. 6.2e-18;
Matches 60; Conservative 26; Mismatches 41; Indels 4; Gaps 3;

QY 1 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
Db 37 SAQNTINATWTQDPSVSGSSVQAMQKGLTNIQLTGHAGVYVSGDGTGSGGLVTIPFK 96
QY 61 DGQGPVFRGRIQGANINDQANTGIDGLA--GWRVASSQETLNVPVTTFGKSTLPAGTF 117
Db 97 NAAQGIPIFRGR-TNADIGQASNTLIAGHSGPGWNLDPDAGNNISLDIKAFQKNDIPAGTY 155
QY 118 TATFYVQYQYQN 128
Db 156 TATFYIQQYQS 166

RESULT 4
Q47405_ECOLI
ID Q47405_ECOLI PRELIMINARY; PRT; 166 AA.
AC Q47405;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Antigen 8786.
GN Name=nfaA;
OS Escherichia coli.
OG Plasmid.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
DR NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8786;
RA Aubel D., Darfeuille-Michaud A., Martin C., Joly B.;
RT "Nucleotide sequence of the nfaA gene encoding the antigen 8786
RT adhesive factor of enterotoxigenic escherichia coli.";
RL FEMS Microbiol. Lett. 98:277-284(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8786;
RX MEDLINE=93093418; PubMed=1281130; DOI=10.1016/0378-1097(92)90169-O;
RA Aubel D., Darfeuille-Michaud A., Martin C., Joly B.;

```


OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RN MEDLINE=97304677; PubMed=9161040;
 RX Serikawa K.A., Martinez-Laborda A., Kim H.S., Zambryski P.C.;
 RA "Localization of expression of KNAT3, a class 2 knotted-like gene."
 RT Plant J. 11:853-861(1997).
 RL
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714; DOI=10.1038/35048507;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuura A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shimpou S., Takeuchi C., Wada T.,
 RA Natanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Balter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Strowmatt C.,
 RA Wagner-McPherson C., Mollam A., Yoakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
 RA Martensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Dusterhoeft A., Stiekema W., Pohl T.,
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
 RA Ramsperger U., Wedler E., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Birke G., Moeljan P., Klein Lankhorst R.,
 RA Weitzenecker T., Dörke W., Rose M., Hauf J., Bernerstorfer S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana."
 RL Nature 408:823-826(2000).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: Belongs to the TALE/KNOX homeobox family.
 CC -1- SIMILARITY: Contains 1 ELK domain.
 CC -1- SIMILARITY: Contains 1 homeobox DNA-binding domain.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; X92392; CAA63130.1; -; mRNA.
 CC EMBL; AC006259; AAC98441.1; -; Genomic_DNA.
 CC TRANSFAC; T04044; -;
 CC GeneFarm; 4056; 408.
 CC InterPro; IPR005539; ELK.
 CC InterPro; IPR001356; Homeobox.
 CC InterPro; IPR012287; Homeobox-rel.
 CC InterPro; IPR005540; KNOX1.
 CC InterPro; IPR005541; KNOX2.
 CC Pfam; PF03789; ELK; 1.
 CC Pfam; PF03790; KNOX1; 1.
 CC Pfam; PF03791; KNOX2; 1.
 CC ProDom; PD000010; Homeobox; 1.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC DNA-binding; Homeobox; Nuclear protein.
 KW DOMAIN 320 343 Homeobox; TALE-type.
 FT DNA_BIND 344 406 Homeobox; 1.
 FT COMPBIAS 23 33 Poly-Pro.
 FT COMPBIAS 34 37 Poly-Gln.
 FT COMPBIAS 58 61 Poly-Asn.

FT COMPBIAS 84 90 Poly-Ala.
 FT COMPBIAS 104 108 Poly-Asn.
 SQ SEQUENCE 431 AA; 47600 MW; 5222B67AB54B9673 CRC64;
 Query Match 14.1%; Score 93.5; DB 1; Length 431;
 Best Local Similarity 33.3%; Pred. No. 6.9;
 Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;
 Qy 1 AAQNTTSANWSQDPGFTGPAVAAGQKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
 Db ATANTTTASSSDSPSSAAAAAANQWLSRSSFLQRNNNNNASIVGDGDDVTGGADTMI 127
 Qy 59 FVDCGQGPVFRGRIQGANINDQAN-TGIDGLAGWRVASSQ-ETLNVPV 104
 Db 128 ----QGEKTKG---GENKNDGGGATADGVVSMQNAHKABILSHPL 168
 RESULT 8
 ID Q8LDC1_ARATH PRELIMINARY; PRT; 431 AA.
 AC Q8LDC1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DS KNAT3 homeodomain protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22089475; PubMed=12093376;
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation."
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
 RN NUCLEOTIDE SEQUENCE.
 RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC EMBL; AY086091; AAM63298.1; -; mRNA.
 CC GO; GO:0005634; C:nucleus; IEA.
 CC GO; GO:0003700; P:transcription factor activity; IEA.
 CC GO; GO:0003355; P:regulation of transcription, DNA-dependent; IEA.
 CC InterPro; IPR005539; ELK.
 CC InterPro; IPR012287; Homeobox-rel.
 CC InterPro; IPR005540; KNOX1.
 CC InterPro; IPR005541; KNOX2.
 CC Pfam; PF03789; ELK; 1.
 CC Pfam; PF03790; KNOX1; 1.
 CC Pfam; PF03791; KNOX2; 1.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX_1; UNKNOWN_1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC DNA-binding; Homeobox; Nuclear protein.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 431 AA; 47557 MW; 5EF832F3CEB2E02B CRC64;
 Query Match 14.1%; Score 93.5; DB 2; Length 431;
 Best Local Similarity 33.3%; Pred. No. 6.9;
 Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;
 Qy 1 AAQNTTSANWSQDPGFTGPAVAAGQKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
 Db ATANTTTASSSDSPSSAAAAAANQWLSRSSFLQRNNNNNASIVGDGDDVTGGADTMI 127
 Qy 59 FVDCGQGPVFRGRIQGANINDQAN-TGIDGLAGWRVASSQ-ETLNVPV 104

Db 128 -----QGENKTTGG-----GENKNDGGGATAADGVVSWQNHAKBILSHPL 168

RESULT 9

ID Q6VA08_TREPA PRELIMINARY; PRT; 398 AA.
AC Q6VA08;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE TprK (Fragment).
GN Name=tprK;
OS Treponema pallidum subsp. pallidum (syphilis treponeme).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=161;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sea81-4;
RX MEDLINE=22925844; PubMed=14563860;
DOI=10.1128/JB.185.21.6262-6268.2003;
RA LaFond R.E., Centurion-Lara A., Godornes C., Rompalo A.M.,
Van Voorhis W.C., Lukehart S.A.;
RT "Sequence diversity of Treponema pallidum subsp. pallidum tprK in
human syphilis lesions and rabbit-propagated isolates.";
RL J. Bacteriol. 185:6262-6268(2003).
DR EMBL; AY346070; AAQ23285.1; -; Genomic DNA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003872; MOSP_C.
DR InterPro; IPR003857; MOSP_Nterm.
DR Pfam; PF02722; MOSP_C; 1.
DR Pfam; PF02707; MOSP_N; 1.
FT NON_TER 1 398
FT NON_TER 398 398
SQ SEQUENCE 398 AA; 43192 MW; C3E55C5467ADEBDB CRC64;

Query Match 14.0%; Score 93; DB 2; Length 398;
Best Local Similarity 30.9%; Pred. No. 7;
Matches 34; Conservative 12; Mismatches 44; Indels 20; Gaps 5;

Qy 13 DPGFTGPAVAAGQKVGTLTITATGPHNSVSIAGKASVSGVATVPFVDDGGQGVPRGRI 72
Db 126 EPGFEG-----AGGKLG-----YKQTDIAGTGLTFDIAFKFASNTDWEKGP--NGNV 170

Qy 73 QGANINDQANTGIDGLAGW---RVASSQETLNVPTTFFGKSTLPAGTFTA 119
Db 171 QAGANHSKYGLGGDILFGWERTREDGVQYIKVELT--GNSTLSSGYATA 218

RESULT 10

ID Q6VA13_TREPA PRELIMINARY; PRT; 400 AA.
AC Q6VA13;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE TprK (Fragment).
GN Name=tprK;
OS Treponema pallidum subsp. pallidum (syphilis treponeme).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=161;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sea81-4;
RX MEDLINE=22925844; PubMed=14563860;
DOI=10.1128/JB.185.21.6262-6268.2003;
RA LaFond R.E., Centurion-Lara A., Godornes C., Rompalo A.M.,
Van Voorhis W.C., Lukehart S.A.;
RT "Sequence diversity of Treponema pallidum subsp. pallidum tprK in
human syphilis lesions and rabbit-propagated isolates.";
RL J. Bacteriol. 185:6262-6268(2003).
DR EMBL; AY346065; AAQ23280.1; -; Genomic DNA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003872; MOSP_C.

DR InterPro; IPR003857; MOSP_Nterm.

DR Pfam; PF02722; MOSP_C; 1.

DR Pfam; PF02707; MOSP_N; 1.

FT NON_TER 1 400

FT NON_TER 400 400

SQ SEQUENCE 400 AA; 43341 MW; AFC15CED22D44400 CRC64;

Query Match 14.0%; Score 93; DB 2; Length 400;
Best Local Similarity 30.9%; Pred. No. 7.1;
Matches 34; Conservative 12; Mismatches 44; Indels 20; Gaps 5;

Qy 13 DPGFTGPAVAAGQKVGTLTITATGPHNSVSIAGKASVSGVATVPFVDDGGQGVPRGRI 72

Db 126 EPGFEG-----AGGKLG-----YKQTDIAGTGLTFDIAFKFASNTDWEKGP--NGNV 170

Qy 73 QGANINDQANTGIDGLAGW---RVASSQETLNVPTTFFGKSTLPAGTFTA 119

Db 171 QAGANHSKYGLGGDILFGWERTREDGVQYIKVELT--GNSTLSSGYATA 218

RESULT 11

ID Q6VA23_TREPA PRELIMINARY; PRT; 401 AA.
AC Q6VA23;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE TprK (Fragment).
GN Name=tprK;
OS Treponema pallidum subsp. pallidum (syphilis treponeme).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=161;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sea81-4;
RX MEDLINE=22925844; PubMed=14563860;
DOI=10.1128/JB.185.21.6262-6268.2003;
RA LaFond R.E., Centurion-Lara A., Godornes C., Rompalo A.M.,
Van Voorhis W.C., Lukehart S.A.;
RT "Sequence diversity of Treponema pallidum subsp. pallidum tprK in
human syphilis lesions and rabbit-propagated isolates.";
RL J. Bacteriol. 185:6262-6268(2003).
DR EMBL; AY346055; AAQ23270.1; -; Genomic DNA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003872; MOSP_C.
DR InterPro; IPR003857; MOSP_Nterm.
DR Pfam; PF02722; MOSP_C; 1.
DR Pfam; PF02707; MOSP_N; 1.
FT NON_TER 1 401
FT NON_TER 401 401
SQ SEQUENCE 401 AA; 43276 MW; 2F9AB2730EC87499 CRC64;

Query Match 14.0%; Score 93; DB 2; Length 401;
Best Local Similarity 30.9%; Pred. No. 7.1;
Matches 34; Conservative 12; Mismatches 44; Indels 20; Gaps 5;

Qy 13 DPGFTGPAVAAGQKVGTLTITATGPHNSVSIAGKASVSGVATVPFVDDGGQGVPRGRI 72

Db 126 EPGFEG-----AGGKLG-----YKQTDIAGTGLTFDIAFKFASNTDWEKGP--NGNV 170

Qy 73 QGANINDQANTGIDGLAGW---RVASSQETLNVPTTFFGKSTLPAGTFTA 119

Db 171 QAGANHSKYGLGGDILFGWERTREDGVQYIKVELT--GNSTLSSGYATA 218

RESULT 12

ID Q6VA22_TREPA PRELIMINARY; PRT; 405 AA.
AC Q6VA22;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE TprK (Fragment).


```

GN Name=tpkK;
OS Treponema pallidum subsp. pallidum (syphilis treponeme).
OC Bacteria; Spirochaetes, Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=161;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sea81-4;
RX MEDLINE=22925844; PubMed=14563860;
RY DOI=10.1128/JB.185.21.6262-6268.2003;
RA LaFond R.E., Centurion-Lara A., Godornes C., Rompalo A.M.,
  van Voorhis W.C., Lukehart S.A.;
RT "Sequence diversity of Treponema pallidum subsp. pallidum tprK in
  human syphilis lesions and rabbit-propagated isolates.";
RL J. Bacteriol. 185:6262-6268(2003).
DR EMBL: AY346056; AAQ23271.1; -; Genomic DNA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR003872; MOSP_C.
DR InterPro: IPR003857; MOSP_Nterm.
DR Pfam: PF02722; MOSP_C; 1.
DR Pfam: PF02707; MOSP_N; 1.
DR NON_TER 1
DR NON_TER 405
DR NON_TER 405
SQ SEQUENCE 405 AA; 43967 MW; AA8B48BA8B12D5E CRC64;

Query Match 14.0%; Score 93; DB 2; Length 405;
Best Local Similarity 30.9%; Pred. No. 7.2;
Matches 34; Conservative 12; Mismatches 44; Indels 20; Gaps 5;

Qy 13 DPGFTGPAVAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFDGQGVFRGRI 72
Db 126 EPQPEG---AGGKLG-----YKQTDIAGTGLTFDIAKFPASNTDMEGKP--NGWV 170

Qy 73 QGANINQANTGIDGLAGW-----RVASSQETLNVPVTFGKSTLPAGTFTA 119
Db 171 QAGANHSKYGIGDILFGWERTRENGVQYIKVELT--GNSLTSSGYATA 218

RESULT 13
Q912M3_PSEAE
ID Q912M3_PSEAE PRELIMINARY; PRT; 2468 AA.
AC Q912M3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA1874;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
  Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
  Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
  Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
  Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
  Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
  opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004613; AAG05263.1; -; Genomic DNA.
DR PIR: A83412; A83412.
DR HSP; P22629.1SWH.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR InterPro: IPR001343; Hemlyen_Ca_bind.
DR InterPro: IPR006162; Ppantne_S.
DR InterPro: IPR002035; VWF_A.
DR PRINTS; PR00313; CARNDRGPT.
DR PRINTS; PR00453; VWFADOMAIN.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.

```

```

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 2468 AA; 238414 MW; 13596AFAB2C4B899 CRC64;

Query Match 13.8%; Score 91.5; DB 2; Length 2468;
Best Local Similarity 26.3%; Pred. No. 69;
Matches 36; Conservative 11; Mismatches 43; Indels 47; Gaps 7;

Qy 7 SANWS-----QDP-GFTGPAVAGQKVGTLISITATGPHNSVSIAGKGA 48
Db 1284 SGWSPTPGTPLANGTVVNAVAQDPAGNTGQ-----GSTTDAVAPNTFVVPNSGN 1336

Qy 49 SVSGGV---ATVPFVFGQGVFRGRIQGANINDQANTGIDGLAGRWVASSQETLN---V 102
Db 1337 LLNGTAEPGSTWLTLDGNGNPI-----GQTTADSGNWSFTPGSQLPNGTVV 1383

Qy 103 PVT---TFGKSTLPAGT 116
Db 1384 NVTASDAAGNTSLPATT 1400

RESULT 14
Q7NP45_GLOVI
ID Q7NP45_GLOVI PRELIMINARY; PRT; 1973 AA.
AC Q7NP45;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE G110213 protein.
GN OrderedLocusNames=g110213;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteriales; Gloeobacteriales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Sakamuro Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
  Nakamura S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
  Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,
  Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
  cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL: BA000045; BAC88154.1; -; Genomic DNA.
DR GO: GO:0016539; F:intron-mediated protein splicing; IEA.
DR InterPro: IPR006141; Intein S.
DR InterPro: IPR002909; IPT/TIG_recept.
DR InterPro: IPR002345; Lipocalin.
DR InterPro: IPR006530; YD.
DR Pfam: PF05593; RHS_repeat; 9.
DR Pfam: PF01833; TIG; 4.
DR TIGRPFAMs; TIGR01443; Intein Cterm; 1.
DR TIGRPFAMs; TIGR01643; YD_repeat_2x; 7.
DR PROSITE; PS00817; INTEIN_NTER; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR Complete proteome.
SQ SEQUENCE 1973 AA; 210549 MW; 9DE8BA267A9B148A CRC64;

Query Match 13.7%; Score 90.5; DB 2; Length 1973;
Best Local Similarity 27.7%; Pred. No. 66;
Matches 39; Conservative 12; Mismatches 45; Indels 45; Gaps 6;

Qy 16 FTGP-----AVAQKQVGTLS-ITATGP-----HNSVSIAGKASVSGGVATVPFDGQG 64
Db 1506 FTGASSVRFNATAAGSVVSAQISATVPTGTAGTGAVSVTTAGGTASGG-----FFDVQA 1560

Qy 65 QPVFRGR-----IQGANINDQANTGIDGLAGRWVASSQETLNVP 103
Db 1561 TPTINSSGVGNPGSGVLSASASDPTVLINGTLNLTNSVTINGIAAYAVSATQIQV- 1619

Qy 104 VTTFGKSTLPAGTFTATFYVQ 124
Db 1620 -----SLPENATGTLVQV 1633

```

Search completed: April 26, 2006, 17:04:38
Job time : 77 secs

;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTION
;; TITLE OF INVENTION: OP SALMONELLA
;; NUMBER OF SEQUENCES: 61
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Seed and Berry
;; STREET: 6300 Columbia Center, 701 Fifth Avenue
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: U.S.A.
;; ZIP: 98104-7092
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; FILING DATE: 26-APR-1994
;; APPLICATION NUMBER: US/08/233,788A
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: King, Joshua
;; REGISTRATION NUMBER: 35,570
;; REFERENCE/DOCKET NUMBER: 920043.403C2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; TELEX: 3723836 SEEDANBERRY
;; INFORMATION FOR SEQ ID NO: 40:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 165 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-233-788A-40

Query Match 100.0%; Score 662; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 8.9e-62;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGSVSGGVATVPFV 60
Db 38 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGSVSGGVATVPFV 97

Qy 61 DGGQPVPRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPTTFGKSTLPAGTFTAT 120
Db 98 DGGQPVPRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPTTFGKSTLPAGTFTAT 157

Qy 121 FVYQQYQN 128
Db 158 FVYQQYQN 165

RESULT 3
US-08-449-922-1
; Sequence 1, Application US/08449922
; Patent No. 5510241
; GENERAL INFORMATION:
; APPLICANT: THORNS, CHRISTOPHER J
; TITLE OF INVENTION: METHOD OF TESTING FOR SALMONELLA
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P C
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/449,922
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08,030,208
;; FILING DATE: 26-MAR-1993
;; APPLICATION NUMBER: GB 9021290.3
;; FILING DATE: 01-OCT-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9022570.7
;; FILING DATE: 17-OCT-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9106546.6
;; FILING DATE: 27-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: CRAWFORD, ARTHUR R
;; REGISTRATION NUMBER: 25,327
;; REFERENCE/DOCKET NUMBER: 1498-30
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 176 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: Salmonella enteritidis/Salmonella dublin
;; US-08-449-922-1

Query Match 100.0%; Score 662; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 9.6e-62;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGSVSGGVATVPFV 60
Db 49 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGSVSGGVATVPFV 108

Qy 61 DGGQPVPRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPTTFGKSTLPAGTFTAT 120
Db 109 DGGQPVPRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPTTFGKSTLPAGTFTAT 168

Qy 121 FVYQQYQN 128
Db 169 FVYQQYQN 176

RESULT 4
US-09-230-078A-4
; Sequence 4, Application US/09230078A
; Patent No. 6495334
; GENERAL INFORMATION:
; APPLICANT: Rajashekara, Gireesh
; APPLICANT: Kakambi, Nagarata V.
; APPLICANT: Kapur, Vivek
; TITLE OF INVENTION: RECOMBINANT SEF14 FIMBRIAL PROTEIN FROM SALMONELLA
; FILE REFERENCE: 600.33SUSNO
; CURRENT APPLICATION NUMBER: US/09/230,078A
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: PCT/US97/12639
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: 60/022,191
; PRIOR FILING DATE: 1996-07-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Salmonella enteritidis
; US-09-230-078A-4

Query Match 100.0%; Score 662; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 9.9e-62;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLSTATGPHNSVSIAGKASVSGGVATVPFV 60
Db 53 AAQNTTSANWSQDPGFTGPAVAAGKVGTLSTATGPHNSVSIAGKASVSGGVATVPFV 112
Qy 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTTAT 120
Db 113 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTTAT 172
Qy 121 FYVQYQYQY 128
Db 173 FYVQYQYQY 180

RESULT 5
US-09-543-407-46
; Sequence 46, Application US/09543407
; Patent No. 6864365
; GENERAL INFORMATION:
; APPLICANT: White, Aaron P.
; APPLICANT: Doran, James L.
; APPLICANT: Collinson, S. Karen
; APPLICANT: Kay, William W.
; TITLE OF INVENTION: BACTERIAL FIBRILLAR SYSTEM FOR
; FILE REFERENCE: 920043.406
; PRESENTATION OF HETEROLOGOUS PEPTIDE SEQUENCES
; CURRENT FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Salmonella enteritidis
US-09-543-407-46

Query Match 97.7%; Score 646.5; DB 2; Length 143;
Best Local Similarity 99.2%; Pred. No. 3.1e-60;
Matches 127; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLSTATGPHNSVSIAGKASVSGGVATVPFV 60
Db 17 AAQNTTSANWSQDPGFTGPAVAAGKVGTLSTATGPHNSVSIAGKASVSGGVATVPFV 76
Qy 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTTAT 120
Db 77 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAG-FTAT 135
Qy 121 FYVQYQYQY 128
Db 136 FYVQYQYQY 143

RESULT 6
US-08-229-287-4
; Sequence 4, Application US/08229287
; Patent No. 5530193
; GENERAL INFORMATION:
; APPLICANT: Clark Jr., John M.
; APPLICANT: Jilka, Joseph M.
; APPLICANT: Murry, Lynn E.
; APPLICANT: Scarafia, Liliana E.
; TITLE OF INVENTION: VIRUS RESISTANT CORN PLANTS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sandoz Agro, Inc.
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,287
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/067,257
; FILING DATE: 25-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/817,922
; FILING DATE: 08-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Marcus-Wyner, Lynn
; REGISTRATION NUMBER: 34,869
; REFERENCE/DOCKET NUMBER: 135-1084/XCC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/354-3588
; TELEFAX: 415/857-1125
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-229-287-4

Query Match 13.2%; Score 87.5; DB 1; Length 328;
Best Local Similarity 30.1%; Pred. No. 0.41;
Matches 31; Conservative 9; Mismatches 54; Indels 9; Gaps 2;
Qy 11 SQDPGFTGPAVAAGKVGTLSTATGPHNSVSIAGKASVSGGVATVPFVDGQGPVFRG 70
Db 12 SCSQGTTPATGSGAK-----PATSGAGSGGTGAGTGTGQARTSGTGTSGATGG 65
Qy 71 RIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLP 113
Db 56 QSGSGSGTGEQVNT---GSAGTNATGQDRDRVDAGSTGKISVP 105

RESULT 7
US-09-252-991A-30227
; Sequence 30227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30227
; LENGTH: 2736
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30227

Query Match 13.1%; Score 86.5; DB 2; Length 2736;
Best Local Similarity 22.8%; Pred. No. 7.9;
Matches 33; Conservative 12; Mismatches 43; Indels 57; Gaps 6;
Qy 7 SANW-----SQDP-GFTGPAVAAGKVGTLSTATGPHNSVSIAGKA 48
Db 1306 SGNWTFPTPATPLANGTVVNAVAQDPAGNTGQ-----GSTTVDVAEPTPVNPSNGN 1358
Qy 49 SVSGSV---ATVPFVDGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLN--- 101
Db 1359 LLNGTAEPSGTVTTLTDGNGNFI-----GQTADSGNWSFTFGSQLPNGTVV 1405

QY 102 -----VPVTFKSTLPA 114
DB 1406 NVTASDAAGNTSAPATTTVDSSLPS 1430

RESULT 8
US-09-712-363-190
; Sequence 190, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 190
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-190

Query Match 12.5%; Score 83; DB 2; Length 549;
Best Local Similarity 25.6%; Pred. No. 2.4;
Matches 34; Conservative 18; Mismatches 43; Indels 38; Gaps 5;

QY 1 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGK-----GASVS- 51
DB 142 AAEKDGAGDPDPDPDPDRDPAALAA--LGTPALAAPAPHGALAGSGKLGVRDVLFGKVS 199
QY 52 ---GGVATVPFDGQGVFGRIGQANINDQANTGIDGLAGWRVASSQBE---TLNVPVT 105
DB 200 LALGILVAIALVIG-----GIGGVIGRKTAEVVDVDAFTTSKVTLS 238
QY 106 TFGKSTLPAGTFT 118
DB 239 TTGNAQBPAGRFT 251

RESULT 9
US-09-252-991A-26438
; Sequence 26438, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26438
; LENGTH: 1415
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26438

Query Match 12.5%; Score 82.5; DB 2; Length 1415;
Best Local Similarity 28.8%; Pred. No. 8.9;
Matches 34; Conservative 13; Mismatches 52; Indels 19; Gaps 6;

QY 2 AONTTSANWSQDPGFT-----GPAV--AAQKVGTLISITATGPHN-----SVSTAGKG 47
DB 392 AQPAGANASDSNGVTIVQQQPAVDLAAGAN-GTSAVQSQSGANIGSGANGISVVQSQNG 450
QY 48 ASVSGGVATVPFDGQGVFGRIGQANI---NDQANTGIDGLAGWRVASSQETLNV 102
DB 451 ANIGAGASDISVVQSQNSPNTGSGVNGVTVVQSQNGANIG-SGASGITVVQSQNGANI 507

RESULT 10
US-08-496-944-2
; Sequence 2, Application US/08496944
; Patent No. 6040496
; GENERAL INFORMATION:
; APPLICANT: Law, Marcus D
; APPLICANT: Dietz, Jon M
; TITLE OF INVENTION: Use of Translationally altered RNA to
; TITLE OF INVENTION: Confer Resistance to Maize Dwarf Mosaic Virus and Other
; TITLE OF INVENTION: Monocotyledonous Plant Viruses
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,944
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1814
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2763 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-496-944-2

Query Match 12.5%; Score 82.5; DB 2; Length 2763;
Best Local Similarity 29.1%; Pred. No. 21;
Matches 30; Conservative 9; Mismatches 55; Indels 9; Gaps 2;

QY 11 SQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKASVSGGVATVPFDVGQGPVPRG 70
DB 2447 SGSQGTTPPATGSGAK-----PATSGAGSGSDTGAGTGTGTSQARTGSGTGTSGATGG 2500
QY 71 RIQGANINDQANTGIDGLAGWRVASSQETLNVPTTFKSTLP 113

QY 102 -----VPVTFKSTLPA 114
DB 1406 NVTASDAAGNTSAPATTTVDSSLPS 1430

RESULT 8
US-09-712-363-190
; Sequence 190, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 190
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-190

Query Match 12.5%; Score 83; DB 2; Length 549;
Best Local Similarity 25.6%; Pred. No. 2.4;
Matches 34; Conservative 18; Mismatches 43; Indels 38; Gaps 5;

QY 1 AAQNTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGK-----GASVS- 51
DB 142 AAEKDGAGDPDPDPDPDRDPAALAA--LGTPALAAPAPHGALAGSGKLGVRDVLFGKVS 199
QY 52 ---GGVATVPFDGQGVFGRIGQANINDQANTGIDGLAGWRVASSQBE---TLNVPVT 105
DB 200 LALGILVAIALVIG-----GIGGVIGRKTAEVVDVDAFTTSKVTLS 238
QY 106 TFGKSTLPAGTFT 118
DB 239 TTGNAQBPAGRFT 251

RESULT 9
US-09-252-991A-26438
; Sequence 26438, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

Db 2501 QSGSGTQVNT---GSGATNATGQRDRDNDAGSTGKISVP 2540

RESULT 11

US-09-252-991A-19167
; Sequence 19167, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19167
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19167

Query Match 12.3%; Score 81.5; DB 2; Length 693;
Best Local Similarity 28.3%; Pred. No. 4.6;
Matches 32; Conservative 17; Mismatches 51; Indels 13; Gaps 6;

QY 15 GFTGPAVAAGQKVGTLSTATGPHNSVSIAGKAS--VSGGVATVPFVDGQG-QPVFRGR 71

Db 236 GLFGRVGGPAKVSFSSALMG-----ITGSGVANVVTGQFTIPLMKRFGYKPAFAGG 280

QY 72 IQG-ANINDQANTGIDGLAGWRVASSQETLNPVTFPGKSTL-PAGTFTATFY 122

Db 281 VEATASMSQSLMPPVMGAVAFIMA---ETINPVYVYIAKAALIPALLYRGSVY 330

RESULT 12

US-09-676-519-27
; Sequence 27, Application US/09676519
; Patent No. 6737508
; GENERAL INFORMATION:
; APPLICANT: PELLETIER, JERRY
; APPLICANT: GROS, PHILIPPE
; APPLICANT: DUBOW, MICHAEL
; TITLE OF INVENTION: DNA SEQUENCES FROM STAPHYLOCOCCUS AUREUS BACTERIOPHAGES
; TITLE OF INVENTION: 3A, 77, AND 96 THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES
; FILE REFERENCE: 073406-0404
; CURRENT APPLICATION NUMBER: US/09/676,519
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/407,804
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/110,992
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 27
; LENGTH: 1509
; TYPE: PRT
; ORGANISM: Staphylococcus bacteriophage
US-09-676-519-27

Query Match 12.0%; Score 79.5; DB 2; Length 1509;
Best Local Similarity 28.6%; Pred. No. 20;
Matches 28; Conservative 9; Mismatches 34; Indels 27; Gaps 3;

QY 9 NWSQDPGFTGPAVAAGQKVGTLST--ATGPH-----NSVS 42

Db 1194 NFSKSP-SGTWWKPGDVGLTGNTGTFGPHLHFEMRNRGRHFDPEYLRNKKGRLS 1252

QY 43 IAGKASVSGGVATVPFVDGQGQVFRGRIOGANINDQ 80

Db 1253 IGGGATSGSGATYASRVIRQAQSILGGRYKGRWIHQ 1290

RESULT 13

US-08-616-844-40
; Sequence 40, Application US/08616844
; Patent No. 5849578
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,844
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-616-844-40

Query Match 11.9%; Score 79; DB 1; Length 1481;
Best Local Similarity 30.1%; Pred. No. 22;
Matches 44; Conservative 16; Mismatches 54; Indels 32; Gaps 8;

QY 6 TSANWSQ-DPGFTGPAVAAGQKV---GTLSSITA---TGPHNSVSIAGKASVSGGVATV 57

Db 931 TSTNLAQMSPTFTTILKTSQPLMTTPGTLSTASLVTGPIAVQTAGKQLSLTHPELV 990

QY 58 PFVDGQGQ-PVFRGRIOGANINDQANTGIDGLAGWRVASSQETLNPVVT-----TFG 108

Db 991 PQISTEGGISTERNRV---IVD-ATTGLIPLTSTVPTSAKEMTTKLGVTAEYSPASRSLG 1045

QY 109 KSTLPAGTF-----TATFYVQ 124

Db 1046 TSPSPQTTVVSTAEDLAPKSAFVQ 1071

RESULT 14

US-08-599-654-40

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:04:53 ; Search time 64 Seconds
(without alignments)
835.659 Million cell updates/sec

Title: US-10-018-892-3
Perfect score: 662
Sequence: 1 AAQNTTSANWSQDPGFTGPA.....KSTLPAGTFTATFYVQYQN 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	14.2	176	4	US-10-425-115-215192
2	94	14.2	233	4	US-10-425-115-215188
3	93.5	14.1	431	4	US-10-374-780A-1998
4	91.5	13.8	2468	4	US-10-246-330-4
5	91.5	13.8	2468	4	US-10-282-122A-66335
6	88.5	13.4	439	4	US-10-425-114-72872
7	87	13.1	609	4	US-10-156-761-10122
8	83.5	12.6	219	4	US-10-425-114-70608
9	83.5	12.6	230	4	US-10-425-115-269298
10	83	12.5	466	4	US-10-156-761-9303
11	83	12.5	524	4	US-10-282-122A-62373
12	83	12.5	549	3	US-09-712-363-190
13	83	12.5	549	4	US-10-282-122A-64511
14	83	12.5	549	4	US-10-432-934-62
15	83	12.5	660	6	US-11-053-710A-5
16	81.5	12.3	1721	4	US-10-282-122A-62548
17	81.5	12.3	2204	4	US-10-282-122A-64364
18	81.5	12.3	2694	4	US-10-184-644-207
19	81.5	12.3	2694	4	US-10-184-634-207
20	81	12.2	174	4	US-10-425-115-215193
21	81	12.2	329	4	US-10-029-386-32216
22	81	12.2	334	4	US-10-282-122A-47939
23	81	12.2	441	4	US-10-425-115-289415
24	81	12.2	605	4	US-10-094-749-3209
25	81	12.2	607	4	US-10-425-115-269306
26	81	12.2	827	4	US-10-282-122A-47312
27	80.5	12.2	139	4	US-10-437-963-183869

28	80.5	12.2	595	4	US-10-282-122A-47723	Sequence 47723, A
29	80.5	12.2	999	4	US-10-408-765A-2033	Sequence 2033, Ap
30	80.5	12.2	1862	4	US-10-282-122A-49757	Sequence 49757, A
31	80	12.1	176	4	US-10-767-701-38981	Sequence 38981, A
32	80	12.1	176	4	US-10-425-115-366647	Sequence 366647, A
33	80	12.1	205	4	US-10-425-115-269307	Sequence 269307, A
34	80	12.1	291	4	US-10-156-761-13735	Sequence 13735, A
35	79.5	12.0	342	4	US-10-156-761-13299	Sequence 13299, A
36	79.5	12.0	561	4	US-10-156-761-13274	Sequence 13274, A
37	79	11.9	367	4	US-10-282-122A-47588	Sequence 47588, A
38	79	11.9	654	4	US-10-467-490-6	Sequence 6, Appli
39	79	11.9	946	3	US-09-840-746-1	Sequence 1, Appli
40	79	11.9	1331	4	US-10-282-122A-47930	Sequence 47930, A
41	79	11.9	1481	3	US-09-371-900-40	Sequence 40, Appl
42	79	11.9	1481	3	US-09-924-417-60	Sequence 60, Appl
43	79	11.9	1481	4	US-10-186-950-40	Sequence 40, Appl
44	79	11.9	1481	4	US-10-653-872-60	Sequence 60, Appl
45	78.5	11.9	158	4	US-10-767-701-45705	Sequence 45705, A

ALIGNMENTS

RESULT 1
US-10-425-115-215192
; Sequence 215192, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 215192
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_127852C.1.pap
US-10-425-115-215192

Query Match	14.2%;	Score 94;	DB 4;	Length 176;
Best Local Similarity	29.5%;	Pred. No. 0.16;		
Matches	38;	Conservative 14;	Mismatches 41;	Indels 36; Gaps 6;
QY	1	AAQNTTSANWS-----QDPGFTGPAVAGQKVGTLISITAT-----G	36	
Db	55	ASSNTSSTNFGAVVAIDPLTTGPTAACTGVRAGQTYTFADQQTGGLLMVNFVFTAG	114	
QY	37	PHN--SVSTAGKSGSGVATVPFDGQGO--PVRGRIGQANINDOANTGIDGLACGRV	93	
Db	115	EHNGTSLTSLGN--EVLSDVREMSIVGGSGKFRMARGVQAHTIDSGATSG-----T	165	
QY	94	ASSQETLNV 102		
Db	166	TVVQYTVNV 174		

RESULT 2
US-10-425-115-215188
; Sequence 215188, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 215188
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_127849C.1.pep
US-10-425-115-215188

Query Match 14.2%; Score 94; DB 4; Length 233;
Best Local Similarity 29.5%; Pred. No. 0.22; Mismatches 14; Indels 36; Gaps 6;
Matches 36; Conservative 14;

Qy 1 AAQNTTSANWS-----ODPGFTGPAVAAAGKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
Db 112 ASSNTSSTNFGAVVAIDDLTTGTTRAAGTEVGRAQGYTFADQQTGLLMVMNFVFTAG 171

Qy 37 PHN--SVSIAGKGVSGGVATVPFVDGQGO-PVFRGRIQGANINDQANTGIDGLACWRV 93
Db 172 EHGSTLSILGRN-EVLSDVREMSIVGSGKFRMARGVYQAHTIDSGATSG-----T 222

Qy 94 ASSQETLNV 102
Db 223 TVVQYTVNV 231

RESULT 3
US-10-374-780A-1998
; Sequence 1998, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: 2003-02-25
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1998
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G426 Paralogous to G427
US-10-374-780A-1998

Query Match 14.1%; Score 93.5; DB 4; Length 431;
Best Local Similarity 33.3%; Pred. No. 0.54; Mismatches 14; Indels 11; Gaps 6;
Matches 36; Conservative 14;

Qy 1 AAQNTTSANWSQDPGFTGPAVAAAGKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
Db 68 ATANTTTASSSDSPSSAAAAAANQWLSRSSFLQRNNNNASIVGDGDDVTGGADTMI 127

Qy 59 FVDGQGVFRGRIQGANINDQAN-TGIDGLAGWRVASSQ-ETLNV 104
Db 128 ----QEMKTKG---GENKNDGGGATAADGVVSMQNRHKAETLSHPL 168

RESULT 4
US-10-246-330-4
; Sequence 4, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; APPLICANT: Mah, Thien-Fah
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; FILE REFERENCE: 14537-002001
; CURRENT APPLICATION NUMBER: US/10/246,330
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2468
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-246-330-4

Query Match 13.8%; Score 91.5; DB 4; Length 2468;
Best Local Similarity 26.3%; Pred. No. 7.6; Mismatches 43; Indels 47; Gaps 7;
Matches 36; Conservative 11;

Qy 7 SANWS-----QDP-GFTGPAVAAAGKVG-TLSITATGPHNSVSIAGKGA 48
Db 1284 SGNWSFTGTPLANGTVVNAVAQDPAGTGPQ-----GSTTDAVAFNTPVVPSNGN 1336

Qy 49 SVSGGV---ATVPFVDGQGVFRGRIQGANINDQANTGIDGLAGWRVASSQETLN---V 102
Db 1337 LLNGTAEPGSTVTLTDGNGNPI-----GQTTADSGSNWSFTPGSQLPNGTVV 1383

Qy 103 PVT---TFGKSTLPAGT 116
Db 1384 NVTASDAAGNTSLPAT 1400

RESULT 5
US-10-282-122A-66335
; Sequence 66335, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

```

```
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66335
; LENGTH: 2468
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66335

Query Match          13.8%; Score 91.5; DB 4; Length 2468;
Best Local Similarity 26.3%; Pred. No. 7.6;
Matches 36; Conservative 11; Mismatches 43; Indels 47; Gaps 7;

Qy      7 SANWS-----QDP-GPTGPAVAAGQKVGTLTITATGPHNSVSIAGKGA 48
Db      1284  SGNWSFTPGTFLANGTVVNAVAQDPAGNTGQ-----GSTTVDVAVAPTPVNPNSGN 1336

Qy      49 SVSGGV---ATVPFVDGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSOETLN---V 102
Db      1337  LLNGTAERGSIIVTLTDNGENFI-----GQTADGSGNWSFTPGSQLPNGTVV 1383

Qy      103 PVT---TFGKSTLPAGT 116
Db      1384  NVTASDAAGNTSLPATT 1400

RESULT 6
US-10-425-114-72872
; Sequence 72872, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, Steven E.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 72872
```

```
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana columbia
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-ATXLIB327421P4P06_FLI.pep
US-10-425-114-72872

Query Match          13.4%; Score 88.5; DB 4; Length 439;
Best Local Similarity 33.3%; Pred. No. 1.8;
Matches 36; Conservative 13; Mismatches 48; Indels 11; Gaps 6;

Qy      1 AAQNTTSANWSQDPGFTGPAVAAGQKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
Db      88  ATANTTTASSDPPSSAAAAAANQWLSPSSFLQRNNNNASIVGDIIDVDTGGADTMI 147

Qy      59 FVDGQGPVFRGRIQGANINDQAN-TGIDGLAGRWVASSO-ETLNVPV 104
Db      148  ----QCEMTGG---GENKIDGGGATAADGVVSWQNAHKAETLSHPL 188

RESULT 7
US-10-156-761-10122
; Sequence 10122, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10122
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10122

Query Match          13.1%; Score 87; DB 4; Length 609;
Best Local Similarity 25.7%; Pred. No. 3.9;
Matches 35; Conservative 22; Mismatches 49; Indels 30; Gaps 6;

Qy      6 TSANWSQDPGFTGPAVAAGQKVGTLTITATGPHNSVSIAGKGSVSGGVATVPFVDGQGG 65
Db      161  TGAYYADDISLVGPGVDAGOPPAAPGLKAGTVTSTSVALSWSAVPGAT-----GY 211

Qy      66 PVFRGRIQGANINDQANTGIDGLA-----GWRVAS-----SOETLNVPVTTF-----GK 109
Db      212  AIYRDGVKARSVTGTSAT-VTGLSPATAYGFQVAAVNDAGESAKSATVPATTTSAGSGGGS 270

Qy      110 STLPA-----GTETATF 121
Db      271  TDLPAHALVGYLHASF 286

RESULT 8
US-10-425-114-70608
; Sequence 70608, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, Steven E.
; APPLICANT: Screen, Steven E.
```

APPLICANT: Tabaka, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 70608
LENGTH: 219
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3600-056-H10_F11.pep
US-10-425-114-70608

Query Match 12.6%; Score 83.5; DB 4; Length 219;
Best Local Similarity 28.2%; Pred. No. 2.5;
Matches 40; Conservative 16; Mismatches 45; Indels 41; Gaps 6;
QY 1 AAQNTTSANWSQDPGFTGPVAVAGQKVGTTLSITATGPHNSV-----SIAGKGASVSG--- 52
Db 53 AATSLVSAALSGSAASGTGAAA-----SVSAAGAAASVSASLTSVSAAGAAASGSAS 104
QY 53 -GVATVPFDGQGPVPRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPTTFGKST 111
Db 105 LGSASVGSATASASLV-----SVSVASSTGSGSLA---TTTSVSTAGAAVSSFGVST 153
QY 112 LPA-----GTFTAT 120
Db 154 ABAASWAASSTGAAGTWTST 175

RESULT 9
US-10-425-115-269298
Sequence 269298, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 269298
LENGTH: 230
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MFT4577_177198C.1.pep
US-10-425-115-269298

Query Match 12.6%; Score 83.5; DB 4; Length 230;
Best Local Similarity 27.3%; Pred. No. 2.7;
Matches 41; Conservative 16; Mismatches 58; Indels 35; Gaps 5;
QY 1 AAQNTTSANWSQDPGFTGPVAVAGQKVGTTLSITATGPHNSVSIAGKGASVSGVATVPFV 60
Db 50 AAASLVSAALSGSAASGTGAAAASVAPGA-AASVSASLASVSAAGAAASGSASLASVP-A 107
QY 61 DGQGP-----VPRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPTTFGKST 111
Db 108 AGAVAPGSASLGSASVGSATASASLVSVASSTGTGSLA---TTTSVSTAGAAVSSFGVST 164
QY 112 LPA-----GTFTAT 120
Db 165 ABAASWAASSTGAAGTWTST 194

RESULT 10

US-10-156-761-9303
Sequence 9303, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9303
LENGTH: 466
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-9303

Query Match 12.5%; Score 83; DB 4; Length 466;
Best Local Similarity 28.9%; Pred. No. 7.3;
Matches 37; Conservative 11; Mismatches 68; Indels 12; Gaps 5;
QY 1 AAQNTTSANWSQDPGFTGPVAVAGQKVGTTLSITATGPHNSVSIAGKGASVSGVATVPFV 60
Db 62 AGSGVSTANWQYDTGTSTYPGGAHWGTGEV-TMTSSNNVSLDG-----NGNLRITPLR 115
QY 61 DGQGPVPRGRIQGANINDQANTGIDGLAGRWVASSQETLNVPTTFGKSTLPAGTFTAT 120
Db 116 DSAGNWT-SGRIETNRTDFQPPAG---GTLRVESRIQLPNV-TGAAGKGYWPAFWNLGA 169
QY 121 FYVQOYQN 128
Db 170 PYRGNQYN 177

RESULT 11
US-10-122A-122A-62373
Sequence 62373, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykend, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06

Matches 34; Conservative 18; Mismatches 43; Indels 38; Gaps 5;
Qy 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLSTATGPHNSVSIAGK-----GASVS- 51
Db 142 AAERKDGAGDPDPDRDPAALAAA--LGTPLAALAPAPHGALAGSKLGVRDVLFGGKVS 199
Qy 52 ---GGVATVPFVDGQGVFRGRIOGANINDQANTGIDGLAGWRVASSQ---TLNVPT 105
Db 200 LALGILVAIALVIG-----GIGGVIGRKTAEVVDFTTSKVTLS 238
Qy 106 TFGKSTLPACTFT 118
Db 239 TTGNAQEPAGRFT 251

RESULT 14
US-10-432-934-62
; Sequence 62, Application US/10432934
; Publication No. US20040110269A1
; GENERAL INFORMATION:
; APPLICANT: Vipond, Richard
; APPLICANT: Shuttleworth, Helen
; APPLICANT: Ambrose, Emma
; APPLICANT: Minton, Nigel Peter
; TITLE OF INVENTION: Protection against mycobacterial infections
; FILE REFERENCE: 1581.0970001
; CURRENT APPLICATION NUMBER: US/10/432,934
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: PCT/GB01/05250
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: GB0028966.0
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-432-934-62

Query Match 12.5%; Score 83; DB 4; Length 549;
Best Local Similarity 25.6%; Pred. No. 8.9;
Matches 34; Conservative 18; Mismatches 43; Indels 38; Gaps 5;
Qy 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLSTATGPHNSVSIAGK-----GASVS- 51
Db 142 AAERKDGAGDPDPDRDPAALAAA--LGTPLAALAPAPHGALAGSKLGVRDVLFGGKVS 199
Qy 52 ---GGVATVPFVDGQGVFRGRIOGANINDQANTGIDGLAGWRVASSQ---TLNVPT 105
Db 200 LALGILVAIALVIG-----GIGGVIGRKTAEVVDFTTSKVTLS 238
Qy 106 TFGKSTLPACTFT 118
Db 239 TTGNAQEPAGRFT 251

RESULT 15
US-11-053-710A-5
; Sequence 5, Application US/11053710A
; Publication No. US20050204408A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, ANTHONY S
; TITLE OF INVENTION: TROPOELASTIN DERIVATIVES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRIFFITH HACK
; STREET: 168 WALKER STREET
; CITY: NORTH SYDNEY
; STATE: NEW SOUTH WALES
; COUNTRY: AUSTRALIA
; ZIP: 2060

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC7DOS/MS7DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/053,710A
FILING DATE: 08-Feb-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P08117
FILING DATE: 18?JUL?1997
ATTORNEY/AGENT INFORMATION:
NAME: GUMLEY, THOMAS P
REFERENCE/DOCKET NUMBER: 04828ZK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 9957 5944
TELEFAX: 61 2 9957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-11-053-710A-5
Query Match 12.5%; Score 83; DB 6; Length 660;
Best Local Similarity 33.8%; Pred. No. 11;
Matches 27; Conservative 6; Mismatches 33; Indels 14; Gaps 2;
Qy 15 GFTGPAVAAGKVGTLSTATGPHNSVSIAGKASVSG-----GVATVPFVDGQGVFRG 70
Db 164 GYGGGVAAGKAGGYTGTGTGVPQAAAAAATAAKFGAGAGFGAIVGVGGAGVPGVPG 223
Qy 71 RIQGANINDQANTGIDGLAG 90
Db 224 AI-----PGIGGIAG 233
Search completed: April 26, 2006, 17:06:07
Job time : 65 secs



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 185944

TO: Ginny Portner
Location: REM/3B02/3C18
Art Unit: 1645
Thursday, April 27, 2006

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: 571-272-2527

Case Serial Number: 10/018892

Paul.schulwitz@uspto.gov

Search Notes

Examiner Portner,

Please review the attached search results.

If you have any questions or if you would like to refine the search query, please feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
REM-1A65
571-272-2527

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:05:14 ; Search time 13 Seconds
(without alignments)
447.782 Million cell updates/sec

Title: US-10-018-892-3
Perfect score: 662
Sequence: 1 AAQNTTSANWSQDPGFTGPA.....KSTLPAGTFTATFVQVQYON 128

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /SID55/ptodata/2/pubpaa/US08 NEW PUB.pap.*
2: /SID55/ptodata/2/pubpaa/US06 NEW PUB.pap.*
3: /SID55/ptodata/2/pubpaa/US07 NEW PUB.pap.*
4: /SID55/ptodata/2/pubpaa/PCT NEW PUB.pap.*
5: /SID55/ptodata/2/pubpaa/US09 NEW PUB.pap.*
6: /SID55/ptodata/2/pubpaa/US10 NEW PUB.pap.*
7: /SID55/ptodata/2/pubpaa/US11 NEW PUB.pap.*
8: /SID55/ptodata/2/pubpaa/US60 NEW PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.5	14.1	431	7	US-11-096-568A-33761
2	93	14.0	171	7	US-11-096-568A-24841
3	93	14.0	189	7	US-11-096-568A-24840
4	83.5	12.6	346	6	US-10-517-939-170
5	82	12.4	347	6	US-10-517-939-222
6	81.5	12.3	2204	7	US-11-052-554A-134
7	81	12.2	5712	7	US-11-143-980-47
8	80	12.1	176	7	US-11-096-568A-21627
9	80	12.1	208	7	US-11-096-568A-21626
10	79	11.9	654	7	US-11-169-041-186
11	77.5	11.7	591	6	US-10-510-386-22
12	77	11.6	617	7	US-11-072-512-3810
13	76.5	11.6	617	7	US-11-072-512-2915
14	76.5	11.6	757	6	US-10-509-472-2
15	75.5	11.4	1268	7	US-11-052-554A-1
16	75.5	11.4	1571	7	US-11-052-554A-2
17	75	11.3	148	7	US-11-096-568A-32672
18	75	11.3	235	7	US-11-096-568A-32671
19	75	11.3	282	7	US-11-096-568A-32670
20	74.5	11.3	431	7	US-11-052-554A-210
21	74.5	11.3	1345	7	US-11-052-554A-282
22	74	11.2	1049	7	US-11-051-720-1563
23	74	11.2	1049	7	US-11-051-720-1565
24	74	11.2	1521	7	US-11-051-720-1564
25	74	11.2	1526	7	US-11-051-720-1566

26	73.5	11.1	710	7	US-11-045-802-2	Sequence 2, Appli
27	73.5	11.1	710	7	US-11-045-802-19	Sequence 19, Appl
28	73.5	11.1	710	7	US-11-045-802-20	Sequence 20, Appl
29	73.5	11.1	803	7	US-11-188-298-2035	Sequence 2035, Ap
30	72	10.9	347	6	US-10-517-939-254	Sequence 254, App
31	72	10.9	478	7	US-11-096-568A-16409	Sequence 16409, A
32	72	10.9	493	7	US-11-096-568A-16408	Sequence 16408, A
33	72	10.9	542	7	US-11-096-568A-16407	Sequence 16407, A
34	72	10.9	634	6	US-10-632-150-26	Sequence 26, Appl
35	72	10.9	634	7	US-11-073-457-26	Sequence 26, Appl
36	72	10.9	634	7	US-11-073-460-26	Sequence 26, Appl
37	71.5	10.8	228	7	US-11-096-568A-9780	Sequence 9780, Ap
38	70.5	10.6	152	7	US-11-087-099-881	Sequence 881, App
39	70.5	10.6	710	7	US-11-045-802-22	Sequence 22, Appl
40	70.5	10.6	7465	7	US-11-087-099-7521	Sequence 7521, Ap
41	70	10.6	273	6	US-10-821-234-1203	Sequence 1203, Ap
42	70	10.6	360	7	US-11-079-463-10442	Sequence 10442, A
43	70	10.6	710	7	US-11-143-980-40	Sequence 40, Appl
44	70	10.6	949	7	US-11-052-554A-6	Sequence 6, Appli
45	70	10.6	1093	7	US-11-079-463-9529	Sequence 9529, Ap

ALIGNMENTS

RESULT 1
US-11-096-568A-33761
; Sequence 33761, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; TITLE OF INVENTION: Theryby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33761
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(431)
; OTHER INFORMATION: Ceres Seq. ID no. 13604019
US-11-096-568A-33761

Query Match 14.1%; Score 93.5; DB 7; Length 431;
Best Local Similarity 33.3%; Pred No. 0.15;
Matches 36; Conservative 14; Mismatches 47; Indels 11; Gaps 6;
Qy 1 AAQNTTSANWSQDPGFTGPAVAAGQKVG-TLSITATGPHNSVSIAGKGA-SVSGGVATVP 58
Db ATANTTTASSSDSPSAAAAAANQWLSRSSFLQNNNNNASIVGDDIVTGGADTMI 127
Qy 59 FVDCGGQVFRGRQGANINDQAN-TGIDGLAGWRVASSQ-ETLVVPV 104
Db 128 ----QGMKTKG---GENKNDGGGTAADGVVSWQARHKAETLSHPL 168

RESULT 2
US-11-096-568A-24841
; Sequence 24841, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; TITLE OF INVENTION: Theryby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 24841

```

; LENGTH: 171
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(171)
; OTHER INFORMATION: Ceres Seq. ID no. 12462325
; *FEATURE:
; NAME/KEY: misc feature
; LOCATION: (133)..(133)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-24841

Query Match      14.0%; Score 93; DB 7; Length 171;
Best Local Similarity 29.5%; Pred. No. 0.053;
Matches 38; Conservative 14; Mismatches 41; Indels 36; Gaps 6;

QY 1 AAQNTTSANWS-----QDPGFTGPAVAGQKVGTLSTAT-----G 36
DB 50 ASNTSSNFGAVVAIDPLTTGPTRAAGTEVGRAQGYTFADQQTGGLLWMVNFVFTAG 109
QY 37 PHN--SVSIAGKASVSGGVATVPFDGQGO-PVFRGRIQGANINDQANTGIDGLAGWRV 93
DB 110 EHNGSTLSILGRN-EVLSDVREMSXVGGSGKFRMARGVYVQAHTIDSGATSG-----T 160
QY 94 ASSQETLNV 102
DB 161 TVVQYTVNV 169

RESULT 3
US-11-096-568A-24840
; Sequence 24840, Application US/11096568A
; Publication No. US2006048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 24840
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(189)
; OTHER INFORMATION: Ceres Seq. ID no. 12462324
; *FEATURE:
; NAME/KEY: misc feature
; LOCATION: (151)..(151)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-24840

Query Match      14.0%; Score 93; DB 7; Length 189;
Best Local Similarity 29.5%; Pred. No. 0.06;
Matches 38; Conservative 14; Mismatches 41; Indels 36; Gaps 6;

QY 1 AAQNTTSANWS-----QDPGFTGPAVAGQKVGTLSTAT-----G 36
DB 68 ASNTSSNFGAVVAIDPLTTGPTRAAGTEVGRAQGYTFADQQTGGLLWMVNFVFTAG 127
QY 37 PHN--SVSIAGKASVSGGVATVPFDGQGO-PVFRGRIQGANINDQANTGIDGLAGWRV 93
DB 128 EHNGSTLSILGRN-EVLSDVREMSXVGGSGKFRMARGVYVQAHTIDSGATSG-----T 178
QY 94 ASSQETLNV 102
DB 179 TVVQYTVNV 187

```

```

RESULT 4
US-10-517-939-170
; Sequence 170, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)....(24)
US-10-517-939-170

Query Match      12.6%; Score 83.5; DB 6; Length 346;
Best Local Similarity 29.5%; Pred. No. 1.1;
Matches 26; Conservative 12; Mismatches 43; Indels 7; Gaps 2;

QY 25 QKVGTLSITATGPHNSVSIAGKASVSGGVATVPFDGQGO-PVFRGRIQGANINDQANTG 84
DB 205 QSSGSSDITVTGEGGSSSSSGGSGGSKSFTVRA-----RGTVGGENIQLVNN- 258
QY 85 IDGLAGWRVASSQETLNVPTTFTGKSTL 112
DB 259 -QTVASWNLTTSMQYNASTSLSGGITV 285

RESULT 5
US-10-517-939-222
; Sequence 222, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 347
; TYPE: PRT

```

```

; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)....(24)
US-10-517-939-222

Query Match      12.4%; Score 82; DB 6; Length 347;
Best Local Similarity 30.3%; Pred. No. 1.5;
Matches 27; Conservative 13; Mismatches 41; Indels 8; Gaps 3;

Qy 25 QKVGTLISITAT-GPHNSVSIAGKASVSGGVATVPFDGQGVFRGRIQGANINDQANT 83
Db 205 QSSGSSDITVTEGGSSSSSSGGGSSGGSKFTVRA-----RGTVGGENIQLVNN 259
Qy 84 GIDGLAGWRVASSQETLNVPTTCKSTL 112
Db 260 --QTVASWNLTTSMQNYNASTSLSGGITV 286

RESULT 6
US-11-052-554A-134
; Sequence 134, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 134
; LENGTH: 2204
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-134

Query Match      12.3%; Score 81.5; DB 7; Length 2204;
Best Local Similarity 24.8%; Pred. No. 17;
Matches 33; Conservative 18; Mismatches 51; Indels 31; Gaps 7;

Qy 17 TGPAV---AAGOK-----VGTLSITATGPHNSVSIAGKASVS---GGVATVPFDGQGG 65
Db 367 TGDIVLSSIAQRAHFGPIIPNITVVGPTTTVAIGGPNVTAITITGGGAINRIPISIPAA 426
Qy 66 PVFRGRIQGANINDQA---NTGIDGLAG-----MRVASSQETLN--VPVTFGK 109
Db 427 PGE-----GNSTNPSSGFFNTGAGGSGFGNGANGSFWNLASATSGASLLNVGALGS 482
Qy 110 STLPAFTTATFY 122
Db 483 GLANVGTTVSGFY 495

RESULT 7
US-11-143-980-47
; Sequence 47, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haitli, Bradley A.
; APPLICANT: Wagenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry

```

```

; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 47
; LENGTH: 5712
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-11-143-980-47

Query Match      12.2%; Score 81; DB 7; Length 5712;
Best Local Similarity 29.0%; Pred. No. 61;
Matches 29; Conservative 15; Mismatches 52; Indels 4; Gaps 3;

Qy 5 TTSANWSQDPGTGPAVAAQKVGTLIS--ITATGPHNSVSIAGKASVSGGVATVPFY-D 61
Db 2910 TESAEWRQEGRRPRRAGVSAGVGTNAHVILEQAPKHPAGVAAEGRKGRGEPPTVPWVLS 2969
Qy 62 GQGQPVFRGRIQGANINDQANTGIDGL-AGMRVASSQETL 100
Db 2970 GASEAGLRAQIEGLRAFADDNPTLDPADVGNLSLASTRALL 3009

RESULT 8
US-11-096-568A-21627
; Sequence 21627, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21627
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(176)
; OTHER INFORMATION: Ceres Seq. ID no. 12405531
US-11-096-568A-21627

Query Match      12.1%; Score 80; DB 7; Length 176;
Best Local Similarity 28.8%; Pred. No. 1;
Matches 32; Conservative 14; Mismatches 37; Indels 28; Gaps 5;

Qy 1 AAQNTTSANWS-----QDPGFTGPAVAAQKVGTLISITAT-----G 36
Db 55 AASNTSSTSGFAMVAIDDLTTGTPRAAGTGVGRAQGTVPADQQTFGLLMVDFVFTAG 114
Qy 37 PH--NSVSIAGKASVSGGVATVPFDGQGG--PVFRGRIQGANINDQANTG 84
Db 115 EHKGSTLSILGRN-EVLSDVREMSIVGGSGKFRMARGVYQVQHTIDSGFKSG 164

RESULT 9
US-11-096-568A-21626
; Sequence 21626, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide

```

```

; TITLE OF INVENTION: Theryby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21626
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(208)
; OTHER INFORMATION: Ceres Seq. ID no. 12405530
US-11-096-568A-21626

Query Match 12.1%; Score 80; DB 7; Length 208;
Best Local Similarity 28.8%; Pred. No. 1.2;
Matches 32; Conservative 14; Mismatches 37; Indels 28; Gaps 5;

QY 1 AAQNTTSANWS-----QDPGFTGPAVAAQKVGTLISITAT-----G 36
Db 87 AASNTSSTSGAMVAIDDDPLTTGTPTRAAGTEVGRAQGTTFADQQTGGLLMVMDVFVFTAG 146
QY 37 PH--NSVSTAGKASVSGGVATVPFVDCGQ-PVFRGRIQGANINDQANTG 84
Db 147 EHKGSTLSILGRN-EVLSDVREMSIVGGSGKFRMARGYVQAHTIDSGFKSG 196

RESULT 10
US-11-169-041-186
; Sequence 186, Application US/11169041
; Publication No. US20060019284A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 186
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-169-041-186

Query Match 11.9%; Score 79; DB 7; Length 654;
Best Local Similarity 30.1%; Pred. No. 6.5;
Matches 44; Conservative 16; Mismatches 54; Indels 32; Gaps 8;

QY 6 TSANWSQ-DPGFTGPAVAAQKV----GTLISITA---TGPHNSVSIAGKASVSGGVATV 57
Db 104 TSTNLAQMSPTFTTILKTSQPLMTTGTLSSTASLTGPIAVQTAGQLSLTHPEILV 163
QY 58 PFVDCGQG-PVFRGRIQGANINDQANTGIDGLAGRWVASSOETLNVPT-----TFG 108
Db 164 PQISTEGGISTERNRV-----IVD-ATTGLIPLTTSVPTSAKEMTTKLGVTAEYSPASRLIG 218
QY 109 KSTLPACTF-----TATEYVQ 124
Db 219 TSPSPQTTWSTAEADLAPKSATPAVQ 244

RESULT 11
US-10-510-386-22
; Sequence 22, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-22

Query Match 11.7%; Score 77.5; DB 6; Length 591;
Best Local Similarity 24.1%; Pred. No. 8;
Matches 35; Conservative 18; Mismatches 45; Indels 47; Gaps 8;

QY 2 AQTNTS-ANW-----SQDPGFTGPAVAAQKVGTLISITATGPHNSVSIAGKASVS 51
Db 108 AQNASAFSQMNKLIQAALAWNKGLTGKQV---KIAVID-SGISPHEELSIAG-GASMV 161
QY 52 GGAVATVPVDCGQGPVFRGRIQGANINDQANTGIDGLA----- 89
Db 162 G--YTASYRDDNGHGHVAGIIGAKHGR-----GIDGIAPGAQLYAVKALDRNGAGDLKGI 216
QY 90 ----GWRVASSQETLNVPTTFGKS 110
Db 217 LQGDWNSIQHGIDIIINMSLVVSGDS 241

RESULT 12
US-11-072-512-3810
; Sequence 3810, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOMYUKI
; APPLICANT: MASAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3810
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3810
```

```
Query Match      11.6%; Score 77; DB 7; Length 675;
Best Local Similarity 26.0%; Pred. No. 11;
Matches 26; Conservative 17; Mismatches 41; Indels 16; Gaps 4;

Qy 9 NWSQDPG-----FTGPAVAGQKVTLSITATGPHNSVSIAGKASVSGGVATVPFVDG 62
Db 476 SWEADGGIGLLFSGSLVSGSTGRLRLWAGVAVSELRCCKSGSGARSS-----VPM 528

Qy 63 QOQVFRGRIQANDQANTGIDG-LAG--WRVASSQBT 99
Db 529 EHELVDGANVVSASFDDSDVMGVGTGTAGTLNFWVSWAEGT 568
```

```
RESULT 13
US-11-072-512-2915
; Sequence 2915, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2915
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2915
```

```
Query Match      11.6%; Score 76.5; DB 7; Length 617;
Best Local Similarity 35.4%; Pred. No. 11;
Matches 28; Conservative 7; Mismatches 29; Indels 15; Gaps 4;

Qy 14 PGFTGPAVAGQKVTLSITATGPHNSVSIAGKASV--SGGVATVPFVDGQGPVFRGR 71
Db 188 PG--GVAGAAG-KAGYPTGTGVGPQAAAAAATAKAFGAGAGVLPGVGGAGVGPVPGA 244

Qy 72 IQGANINDQANTGIDGLAG 90
Db 245 I-----PGIGGIAG 253
```

```
RESULT 14
US-10-509-472-2
; Sequence 2, Application US/10509472
; Publication No. US20060079439A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: ELASTIN PREVENTS OCCLUSION OF BODY VESSELS BY VASCULAR SMOOTH
; TITLE OF INVENTION: MUSCLE CELLS
```

```
; FILE REFERENCE: HYDR-PWO-005
; CURRENT APPLICATION NUMBER: US/10/509,472
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: 60/368084
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 2
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-509-472-2
```

```
Query Match      11.6%; Score 76.5; DB 6; Length 757;
Best Local Similarity 35.4%; Pred. No. 14;
Matches 28; Conservative 7; Mismatches 29; Indels 15; Gaps 4;

Qy 14 PGFTGPAVAGQKVTLSITATGPHNSVSIAGKASV--SGGVATVPFVDGQGPVFRGR 71
Db 232 PG--GVAGAAG-KAGYPTGTGVGPQAAAAAATAKAFGAGAGVLPGVGGAGVGPVPGA 288

Qy 72 IQGANINDQANTGIDGLAG 90
Db 289 I-----PGIGGIAG 297
```

```
RESULT 15
US-11-052-554A-1
; Sequence 1, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent In version 3.3
; SEQ ID NO 1
; LENGTH: 1268
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-1
```

```
Query Match      11.4%; Score 75.5; DB 7; Length 1268;
Best Local Similarity 31.1%; Pred. No. 32;
Matches 32; Conservative 13; Mismatches 45; Indels 13; Gaps 5;

Qy 22 AAGQKVTLSITATGPHN--SVSIAGKASVSGGVATVPFVDGQ--QVFRGRIQGAN 76
Db 216 AAGQVTG--SVSATGSRNGYSVDVAKLGMANKLSLSTKGVGVNRLGVLAGGVNGVS 273

Qy 77 INDQANTGIDGLAGWRVASSQBTLNVPVTTFGKSTLPAGTFTA 119
Db 274 IDSKGNL----LNSNAIQSASTIN--LTNGTLDNTTGTITS 310
```

Search completed: April 26, 2006, 17:06:26
Job time : 14 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:06:19 ; Search time 186 Seconds
(without alignments)
302.368 Million cell updates/sec

Title: US-10-018-892-3

Perfect score: 128

Sequence: 1 AAQWTSANWSQDFGTGPA.....KSTLPAGTFTATFYVQYQN 128

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2443163 seqs, 439378781 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2441974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq 21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	128	AAB31431	Amino aci
2	128	100.0	144	AAW47082	Salmonell
3	128	100.0	165	AAW23571	Salmonell
4	128	100.0	165	AAB31430	Amino aci
5	128	100.0	176	AAB23731	Fimbrial
6	128	100.0	176	AAW42173	SE fimbri
7	128	100.0	180	AAW47081	Salmonell
8	66	51.6	165	AAW62752	Sefa sequ
9	8	6.2	225	ADJ51007	Human nov
10	8	6.2	345	ADA55453	Human pro
11	8	6.2	490	ADS43323	Bacterial
12	8	6.2	492	ABP69377	Human pol
13	8	6.2	494	ADY37608	lung canc
14	8	6.2	559	ABU62853	Novel end
15	8	6.2	574	ABG29271	Novel hum
16	8	6.2	599	ADC58096	Microtubu
17	8	6.2	705	AAE32112	Human cyt
18	7	5.5	17	AAB57669	CD148 PL
19	7	5.5	17	AAB57692	CD148 PL
20	7	5.5	17	AAB58079	CD148 PL
21	7	5.5	17	AAB58102	CD148 PL
22	7	5.5	17	AAB55830	PL peptid
23	7	5.5	17	AAB55807	PL peptid
24	7	5.5	17	ABJ05199	C-termina

25	7	5.5	17	5	ABJ05222	C-termina
26	7	5.5	17	5	ABP63217	CD148 PL
27	7	5.5	17	5	ABP63247	CD148 PL
28	7	5.5	55	3	AAG05854	Arabidops
29	7	5.5	58	7	ABR61594	Z. mays a
30	7	5.5	63	4	AAW25922	Human pro
31	7	5.5	63	4	AAW79596	Human pro
32	7	5.5	63	4	AAU56367	Propionib
33	7	5.5	63	6	ABM52886	Human rep
34	7	5.5	83	4	AAW95904	Human tes
35	7	5.5	83	4	ABB96435	Human tes
36	7	5.5	84	8	ADS07289	Staphyloc
37	7	5.5	86	3	AAG05852	Arabidops
38	7	5.5	87	4	AAW88979	Human imm
39	7	5.5	95	5	ABP10155	Human ORF
40	7	5.5	104	7	ADD36246	Frog orth
41	7	5.5	104	8	ADQ08869	Xenopus l
42	7	5.5	104	9	ADM46048	African c
43	7	5.5	114	4	AAE12624	Human gen
44	7	5.5	114	4	AAU04354	Mammalian
45	7	5.5	114	4	AAG93282	Human pro

ALIGNMENTS

RESULT 1

AAB31431
ID AAB31431 standard; peptide; 128 AA.

AC AAB31431;

DT 20-APR-2001 (first entry)

XX Amino acid sequence of the C128 fragment of the Sef14 antigen.

XX C128 fragment; fimbrial Sef14 antigen; fimbrial protein;
XX flagellin protein; poultry.

OS Salmonella enteritidis.

XX WO200078995-A1.

PD 28-DEC-2000.

PF 22-JUN-1999; 99WO-SG000061.

PR 22-JUN-1999; 99WO-SG000061.

XX (MOLE-) INST MOLECULAR AGROBIOLOGY.

XX Kwang H, Liu W, Low SS, Loh KYH;

XX WPI; 2001-071400/08.

XX New method for the specific detection of Salmonella enteritidis
XX infections of poultry comprises contacting a biological sample with
XX antigenic fragments of S. enteritidis fimbrial and/or flagellin proteins.

XX Claim 18; Page 42; 49pp; English.

XX The present sequence represents the C128 fragment of the fimbrial Sef14
XX antigen of Salmonella enteritidis. The specification describes a method
XX for detecting S. enteritidis in a biological sample obtained from
XX poultry. The method comprises contacting the sample with an antigenic
XX fragment of S. enteritidis fimbrial or flagellin protein and detecting
XX the formation of a complex, where the fragment is specifically recognized
XX by S. enteritidis antibodies. The antigenic fragments are specific to
XX Salmonella enteritidis and enable specific detection of S. enteritidis
XX even in the presence of other Salmonella spp. The methods are useful for
XX the specific detection of S. enteritidis infections in biological samples
XX derived from poultry

SQ Sequence 128 AA;
 Query Match 100.0%; Score 128; DB 4; Length 128;
 Best Local Similarity 100.0%; Pred. No. 2.4e-119;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAOQNTTSANWSQDPGFTGPAVAGQKVGTLSITATGPHNSVSIAGKGSVSGGVATVPFV 60
 DB 1 AAOQNTTSANWSQDPGFTGPAVAGQKVGTLSITATGPHNSVSIAGKGSVSGGVATVPFV 60
 QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNPVPTTFGKSTLPAGTFTAT 120
 DB 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNPVPTTFGKSTLPAGTFTAT 120
 QY 121 FYVQOYQN 128
 DB 121 FYVQOYQN 128

RESULT 2
 AAW47082
 ID AAW47082 standard; protein; 144 AA.
 XX
 AC AAW47082;
 XX 25-JUN-1998 (first entry)
 DT Salmonella Sef14 protein fragment.
 XX
 DE
 KW Sefa gene; Sef14 protein; infection; bird; chicken; turkey;
 KW anti-Salmonella enteritidis antibody; vaccine; poultry.
 XX
 OS Salmonella enteritidis.
 XX WO9803656-A1.
 PN 29-JAN-1998.
 PD 18-JUL-1997; 97WO-US012639.
 PF 19-JUL-1996; 96US-0022191P.
 PR (MINU) UNIV MINNESOTA.
 PA Rajashekara G, Nagaraja KV, Kapur V;
 PI WPI; 1998-120780/11.
 DR N-PSDB; AAV13948.
 XX
 XX Detecting antibodies against Salmonella enteritidis using truncated
 PT fibrillar antigen Sef14 - in immunoassays, particularly for diagnosing
 PT infection in poultry, also new antigens.
 XX
 PS Claim 4; Page 23-24; 38pp; English.
 XX
 CC This sequence is a fragment of the Salmonella enteritidis (Se) Sef14
 CC protein, encoded by the sefa gene. The method of the invention is for
 CC detecting antibodies (Ab) against Se in an animal by treating a sample
 CC with a truncated Sef14 antigen (Ag), lacking at least the native Sef14
 CC signal peptide, and detecting any Ab-Ag complex formed. Detection (by
 CC enzyme-linked immunosorbent assay or agglutination tests) of the Ab is
 CC used to diagnose Se infection in birds, especially chickens and turkeys.
 CC The Ag can also be used in vaccines to protect poultry against Se
 CC infection. Detection of the Ab is a sensitive, specific method for
 CC reliable and routine screening of animals. The Ag are easily produced in
 CC large quantities, in pure form, without requiring special growing
 CC conditions, so are suitable for large scale screening of flocks
 XX
 SQ Sequence 144 AA;

Query Match 100.0%; Score 128; DB 2; Length 144;
 Best Local Similarity 100.0%; Pred. No. 2.6e-119;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAOQNTTSANWSQDPGFTGPAVAGQKVGTLSITATGPHNSVSIAGKGSVSGGVATVPFV 60
 DB 17 AAOQNTTSANWSQDPGFTGPAVAGQKVGTLSITATGPHNSVSIAGKGSVSGGVATVPFV 76
 QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNPVPTTFGKSTLPAGTFTAT 120
 DB 77 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNPVPTTFGKSTLPAGTFTAT 136
 QY 121 FYVQOYQN 128
 DB 137 FYVQOYQN 144

RESULT 3
 AAW23571
 ID AAW23571 standard; protein; 165 AA.
 XX
 AC AAW23571;
 XX 25-MAR-2003 (revised)
 DT 29-SEP-1997 (first entry)
 XX Salmonella enteritidis sefa.
 DE Enteropathogenic bacteria; enterobacteria; S. enteritidis; antibody.
 KW Salmonella enteritidis.
 OS US5635617-A.
 PN 03-JUN-1997.
 PD 26-APR-1994; 94US-00233788.
 PF 26-APR-1993; 93US-00054452.
 PR (UYVI-) UNIV VICTORIA INNOVATION & DEV CORP.
 PA Collinson SK, Kay WW, Doran JL;
 PI WPI; 1997-309886/28.
 DR N-PSDB; AAT74143.
 XX Isolated Salmonella gene agfa - used for diagnosis of Salmonella or
 PT enteropathogenic bacteria of the Enterobacteria family.
 XX Example 9; Fig 2; 85pp; English.
 XX
 CC The present sequence represents sefa found in the sef gene cluster from
 CC Salmonella enteritidis. The nucleic acid can be used to provide
 CC diagnostic assays for Salmonella and/or enteropathogenic bacteria of the
 CC family Enterobacteria. It can also be used to provide proteins and
 CC antibodies which can be used for assays. The nucleic acid sequence can be
 CC used to provide probes or primers which can specifically hybridise to
 CC nucleic acid molecules from greater than 99% of Salmonella strains that
 CC are pathogenic to warm-blooded animals relative to nucleic acid molecules
 CC from virtually all other microbial organisms. (Updated on 25-MAR-2003 to
 CC correct PF field.)
 XX
 SQ Sequence 165 AA;
 Query Match 100.0%; Score 128; DB 2; Length 165;
 Best Local Similarity 100.0%; Pred. No. 2.9e-119;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAOQNTTSANWSQDPGFTGPAVAGQKVGTLSITATGPHNSVSIAGKGSVSGGVATVPFV 60
 DB 38 AAOQNTTSANWSQDPGFTGPAVAGQKVGTLSITATGPHNSVSIAGKGSVSGGVATVPFV 97
 QY 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNPVPTTFGKSTLPAGTFTAT 120
 DB 98 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNPVPTTFGKSTLPAGTFTAT 157

Qy 121 FYVOQYQN 128
 |||||
 Db 158 FYVOQYQN 165

RESULT 4

AAB31430
 ID AAB31430 standard; protein; 165 AA.

XX AC AAB31430;

XX DT 20-APR-2001 (first entry)

XX DE Amino acid sequence of the Sef14 antigen.

XX KW C128 fragment; fimbrial Sef14 antigen; fimbrial protein;
 flagellin protein; poultry.

XX OS Salmonella enteritidis.

XX PN WO200078995-A1.

XX PD 28-DEC-2000.

XX PF 22-JUN-1999; 99WO-SG000061.

XX PR 22-JUN-1999; 99WO-SG000061.

XX PA (MOLE-) INST MOLECULAR AGROBIOLOGY.

XX PI Kwang H, Liu W, Low SS, Loh KYH;

XX DR WPI; 2001-071400/08.

XX DR N-PSDB; AAF24784.

XX PT New method for the specific detection of Salmonella enteritidis
 infections of poultry comprises contacting a biological sample with
 PT antigenic fragments of S. enteritidis fimbrial and/or flagellin proteins.

XX PS Disclosure; Page 42; 49pp; English.

XX CC The present sequence represents the fimbrial Sef14 antigen of Salmonella
 CC enteritidis. The specification describes a method for detecting S.
 CC enteritidis in a biological sample obtained from poultry. The method
 CC comprises contacting the sample with an antigenic fragment of S.
 CC enteritidis fimbrial or flagellin protein and detecting the formation of
 CC a complex, where the fragment is specifically recognized by S.
 CC enteritidis antibodies. The antigenic fragments are specific to
 CC Salmonella enteritidis and enable specific detection of S. enteritidis
 CC even in the presence of other Salmonella spp. The methods are useful for
 CC the specific detection of S. enteritidis infections in biological samples
 CC derived from poultry

XX SQ Sequence 165 AA;

Query Match 100.0%; Score 128; DB 4; Length 165;

Best Local Similarity 100.0%; Pred. No. 2.9e-119;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANWSQDPGFTGPAVAAGOKVGTLSITATGPHNSVSIAGKASVSGGVATVPFV 60

Db 38 AAQNTTSANWSQDPGFTGPAVAAGOKVGTLSITATGPHNSVSIAGKASVSGGVATVPFV 97

Qy 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120

Db 98 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 157

Qy 121 FYVOQYQN 128

|||

Db 158 FYVOQYQN 165

RESULT 5

AAR23731
 ID AAR23731 standard; protein; 176 AA.

XX AC AAR23731;

XX DT 25-MAR-2003 (revised)

XX DT 02-NOV-1992 (first entry)

XX DE Fimbrial antigen.

XX KW S. dublin; S. moscow; fimbria-like structure; epitope.

XX OS Salmonella enteritidis.

XX PN WO9206198-A.

XX PD 16-APR-1992.

XX PF 01-OCT-1991; 91WO-GB001691.

XX PR 01-OCT-1990; 90GB-00021338.

XX PR 17-OCT-1990; 90GB-00022570.

XX PA (UKAG-) UK MIN AGRIC FISH.

XX PI Woodward MJ;

XX DR WPI; 1992-150883/18.

XX PT Detection and identification of salmonella - by using monoclonal
 antibodies to detect epitope(s) of these serotypes in culture.

XX PS Disclosure; Page 3; 48pp; English.

XX CC The sequence given is Salmonella enteritidis fimbrial antigen (SEFA).
 CC Salmonella organisms have fimbria-like structures on their surfaces and
 CC it has been suggested that there are antigenically distinct types of
 CC fimbria, ie. possessing specific epitopes on the fimbrial antigens. This
 CC sequence has an amino acid sequence which forms an epitope on the fimbria
 CC "in vivo" which is specifically encoded by DNA of the species S.
 CC enteritidis, and some strains of the species S. dublin and S. moscow but
 CC which is apparently absent in virtually all other serotypes. This
 CC antigen can be used for testing for the presence of Salmonella
 CC microorganisms in clinical samples such as animal remains or prods., food
 CC samples and infected environmental samples. (Updated on 25-MAR-2003 to
 CC correct PF field.)

XX SQ Sequence 176 AA;

Query Match 100.0%; Score 128; DB 2; Length 176;

Best Local Similarity 100.0%; Pred. No. 3.1e-119;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANWSQDPGFTGPAVAAGOKVGTLSITATGPHNSVSIAGKASVSGGVATVPFV 60

Db 49 AAQNTTSANWSQDPGFTGPAVAAGOKVGTLSITATGPHNSVSIAGKASVSGGVATVPFV 108

Qy 61 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120

Db 109 DGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 168

Qy 121 FYVOQYQN 128

|||

Db 169 FYVOQYQN 176

RESULT 6

AAR42173
 ID AAR42173 standard; protein; 176 AA.

XX AC AAR42173;

DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 03-MAY-1994 (first entry)
 XX SE fimbrial antigen (SEFA).
 XX Salmonella enteritidis fimbrial antigen; SEFA; Salmonella typhi;
 KW Salmonella dublin; serotype.
 XX Salmonella enteritidis.
 OS Salmonella typhi.
 XX WO9320231-A1.
 XX 14-OCT-1993.
 XX 29-MAR-1993; 93WO-GB000647.
 PF 31-MAR-1992; 92GB-00007069.
 PR (UKAG-) UK MIN FISHERIES & FOOD.
 XX Woodward MJ, Thorns CJ;
 XX WPI; 1993-336937/42.
 DR N-PSDB; AAQ49882.
 XX Testing for Salmonella serotypes, esp. S. Typhi - using test kit for
 PT detecting nucleic acid sequences specific to certain sero-types.
 PT Claim 2; Page 19-21; 37pp; English.
 XX DNA encoding SEFA is common to members of the enteritidis, dublin and
 CC typhi serogroups and can therefore be used in the detection of such
 CC organisms. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-
 CC AUG-2003 to correct OS field.)
 XX Sequence 176 AA;
 Query Match 100.0%; Score 128; DB 2; Length 176;
 Best Local Similarity 100.0%; Pred. No. 3.1e-119;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
 Db 49 AAQNTTSANWSQDPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 108
 QY 61 DGGQOPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
 Db 109 DGGQOPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 168
 QY 121 FVQOYQN 128
 Db 169 FVQOYQN 176
 RESULT 7
 ID AA47081
 XX AA47081 standard; protein; 180 AA.
 AC AA47081;
 XX 25-JUN-1998 (first entry)
 XX Salmonella Sef14 protein fragment.
 DE Sefa gene; Sef14 protein; infection; bird; chicken; turkey;
 KW anti-Salmonella enteritidis antibody; vaccine; poultry.
 XX Salmonella enteritidis.
 OS WO9803656-A1.
 PN 26-APR-1994; 94WO-IB000207.
 XX

PD 29-JAN-1998.
 XX 18-JUL-1997; 97WO-US012639.
 PF 19-JUL-1996; 96US-0022191P.
 XX (MINU) UNIV MINNESOTA.
 XX Rajashekara G, Nagaraja KV, Kapur V;
 PI WPI; 1998-120780/11.
 XX N-PSDB; AAV13974.
 DR Detecting antibodies against Salmonella enteritidis using truncated
 PT fimbrial antigen Sef14 - in immunoassays, particularly for diagnosing
 PT infection in poultry, also new antigens.
 XX Claim 3; Page 21-22; 38pp; English.
 XX This sequence is a fragment of the Salmonella enteritidis (Se) Sef14
 CC protein, encoded by the sefa gene. The method of the invention is for
 CC detecting antibodies (Ab) against Se in an animal by treating a sample
 CC with a truncated Sef14 antigen (Ag), lacking at least the native Sef14
 CC signal peptide, and detecting any Ab-Ag complex formed. Detection (by
 CC enzyme-linked immunosorbent assay or agglutination tests) of the Ab is
 CC used to diagnose Se infection in birds, especially chickens and turkeys.
 CC The Ag can also be used in vaccines to protect poultry against Se
 CC infection. Detection of the Ab is a sensitive, specific method for
 CC reliable and routine screening of animals. The Ag are easily produced in
 CC large quantities, in pure form, without requiring special growing
 CC conditions, so are suitable for large scale screening of flocks
 XX Sequence 180 AA;
 Query Match 100.0%; Score 128; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 3.2e-119;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
 Db 53 AAQNTTSANWSQDPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 112
 QY 61 DGGQOPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
 Db 113 DGGQOPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 172
 QY 121 FVQOYQN 128
 Db 173 FVQOYQN 180
 RESULT 8
 ID AAR62752
 XX AAR62752 standard; protein; 165 AA.
 AC AAR62752;
 XX 21-OCT-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 26-JUN-1995 (first entry)
 XX Sefa sequence.
 DE Salmonella; Sefa; vaccine.
 KW Salmonella sp.
 OS Unidentified.
 XX WO9425598-A2.
 PN 10-NOV-1994.
 PD 26-APR-1994; 94WO-IB000207.
 XX

XX 26-APR-1993; 93US-00054452.
 XX (UYVI-) UNIV VICTORIA INNOVATION & DEV CORP.
 PA (KING/) KING J.
 XX Kay WW, Collinson SK, Clouthier SC, Doran JL;
 PI WPI; 1994-358275/44.
 DR N-PSDB; AAQ73061.
 XX
 PT Eliciting an immune response to Salmonella - using attenuated Salmonella
 PT strains, vector constructs, or compsns. contg. fimbrial type proteins.
 XX
 PS Disclosure; Fig 2A-2D; 95pp; English.
 XX
 CC The isolated Sefa protein may be used in a vaccine composition to elicit
 CC an immune response to Salmonella in animals (e.g. food producing animals)
 CC and humans. (Updated on 25-MAR-2003 to correct PN field.)
 CC
 CC Revised record issued on 21-OCT-2004 : Correction to OS line
 XX
 SQ Sequence 165 AA;
 Query Match 51.6%; Score 66; DB 2; Length 165;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 QGQPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTTATATY 122
 DB 100 QGQPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTTATATY 159
 QY 123 VQYQON 128
 DB 160 VQYQON 165
 RESULT 9
 ADJ51007
 ID ADJ51007 standard; protein; 225 AA.
 AC ADJ51007;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human novel protein NOV30a.
 XX
 KW Human; NOVX; autoimmune disease; Alzheimer's disease; stroke; allergy;
 KW Parkinson's disease; Huntington's disease; multiple sclerosis; anxiety;
 KW pain; diabetes; graft versus host disease; pancreatitis; obesity; ulcer;
 KW anaemia; cancer; viral infection; bacterial infection;
 KW parasitic infection.
 XX
 OS Homo sapiens.
 XX
 XX US2004030096-A1.
 PN
 PD 12-FEB-2004.
 XX
 XX 01-AUG-2002; 2002US-00210281.
 PF
 XX 02-AUG-2001; 2001US-0309501P.
 PR
 PR 03-AUG-2001; 2001US-0310291P.
 PR
 PR 08-AUG-2001; 2001US-0310951P.
 PR
 PR 09-AUG-2001; 2001US-0311292P.
 PR
 PR 13-AUG-2001; 2001US-0311979P.
 PR
 PR 14-AUG-2001; 2001US-0312203P.
 PR
 PR 17-AUG-2001; 2001US-0313201P.
 PR
 PR 20-AUG-2001; 2001US-0313643P.
 PR
 PR 20-AUG-2001; 2001US-0313702P.
 PR
 PR 21-AUG-2001; 2001US-0314031P.
 PR
 PR 23-AUG-2001; 2001US-0314466P.
 PR
 PR 28-AUG-2001; 2001US-0315403P.

PR 29-AUG-2001; 2001US-0315853P.
 PR 05-MAR-2002; 2002US-0361775P.
 PR 05-MAR-2002; 2002US-0361832P.
 XX
 PA (GORM/) GORMAN L.
 PA (ZERH/) ZERHUSEN B D.
 PA (EDIN/) EDINGER S R.
 PA (PADI/) PADIGARU M.
 PA (GUOX/) GUO X.
 PA (KEKU/) KEKUDA R.
 PA (ZHON/) ZHONG M.
 PA (PATT/) PATTURAJAN M.
 PA (MILL/) MILLER C E.
 PA (JIWW/) JI W.
 PA (PENA/) PENA C E A.
 PA (BURG/) BURGESS C E.
 PA (SCIO/) SCIORE P.
 PA (STON/) STONE D J.
 PA (TAUF/) TAUPIER R J.
 PA (CASW/) CASMAN S J.
 PA (ROTH/) ROTHENBERG M E.
 PA (MALY/) MALYANKAR U M.
 PA (BOLD/) BOLDOG F L.
 XX
 Gorman L, Zerhusen BD, Edinger SR, Padigaru M, Guo X, Kékuda R;
 PI Zhong M, Patturajan M, Miller CE, Ji W, Pena CE, Burgess CE;
 PI Sciore P, Stone DJ, Taupier RJ, Casman SJ, Rothenberg ME;
 PI Malyankar UM, Boldog FL;
 XX WPI; 2004-168942/16.
 DR N-PSDB; ADJ51006.
 XX
 PT New NOVX polypeptides and polynucleotides, useful in diagnosing, treating
 PT or preventing diseases or conditions, e.g. autoimmune disease,
 PT Alzheimer's disease, diabetes, graft versus host disease, cancer or viral
 PT or bacterial infections.
 XX
 PS Claim 2; SEQ ID NO 72; 342pp; English.
 XX
 CC The invention relates to an isolated NOVX polypeptide (of 44 disclosed)
 CC comprising its mature form, a sequence having at least 95% sequence
 CC identity to NOVX or a sequence comprising one or more conservative
 CC substitutions in the amino acid sequence of NOVX. Also included are a
 CC composition comprising NOVX and a carrier, a kit comprising, in one or
 CC more containers, the composition, a method of identifying an agent that
 CC binds to NOVX, a method for identifying a potential therapeutic agent for
 CC use in treatment of a pathology related to aberrant expression or
 CC aberrant physiological interactions of NOVX, a method for screening for a
 CC modulator of activity of or of latency or predisposition to a pathology
 CC associated with NOVX, a method for modulating the activity of NOVX, a
 CC method of treating or preventing a pathology associated with NOVX or a
 CC pathological state in a mammal, an isolated nucleic acid molecule
 CC encoding a NOVX protein, a vector comprising the nucleic acid molecule,
 CC a cell comprising the vector, an antibody that immunospecifically binds
 CC to NOVX, a method for determining the presence or amount of NOVX or the
 CC nucleic acid molecule in a sample, a method for determining the presence
 CC of or predisposition to a disease associated with altered levels of
 CC expression of NOVX or the nucleic acid molecule in a first mammalian
 CC subject and a method of producing NOVX (comprising culturing the cell
 CC under conditions that lead to expression of the polypeptide). NOVX is
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease associated with NOVX. The polypeptides
 CC and nucleic acid molecules are useful in diagnosing, treating or
 CC preventing diseases or conditions, e.g. autoimmune disease, Alzheimer's
 CC disease, stroke, allergies, Parkinson's disease, Huntington's disease,
 CC multiple sclerosis, anxiety, pain, diabetes, graft versus host disease,
 CC pancreatitis, obesity, ulcers, anaemia, cancer, viral or bacterial and
 CC parasitic infections (many more diseases and disorders are listed in the
 CC specification). The present sequence represents a NOVX protein.
 XX
 SQ Sequence 225 AA;
 Query Match 6.2%; Score 8; DB 8; Length 225;

Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AGQKVGTL 30
|||||
Db 16 AGQKVGTL 23

RESULT 10
ADA55453
ID ADA55453 standard; protein; 345 AA.
XX
AC ADA55453;
XX
DT 20-NOV-2003 (first entry)
XX
XX Human protein, SEQ ID 3021.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
XX 21-MAR-2002; 2002EP-00006586.
PF
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maeuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA53814.
XX
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 3021; 205pp; English.
XX
XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 345 AA;
Query Match 6.2%; Score 8; DB 6; Length 345;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AGQKVGTL 30
|||||
Db 142 AGQKVGTL 149

RESULT 11
ADS43323
ID ADS43323 standard; protein; 490 AA.
XX
AC ADS43323;
XX
DT 02-DEC-2004 (first entry)

XX Bacterial polypeptide #21753.
DE
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
OS
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
PF
PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
FI WPI; 2004-061375/06.
XX
DR New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
PT
XX Claim 1; SEQ ID NO 21753; 122pp; English.
PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 490 AA;

Query Match 6.2%; Score 8; DB 8; Length 490;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AVAAGQKV 27
|||||
Db 268 AVAAGQKV 275

RESULT 12
ABP69377
ID ABP69377 standard; protein; 492 AA.
XX
AC ABP69377;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polypeptide SEQ ID NO 1424.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoasacide;
KW antiarthritic.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR N-PSDB; ABZ11594.
XX
XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 9; SEQ ID NO 1424; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 492 AA;
Query Match 6.2%; Score 8; DB 5; Length 492;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 23 AGQKVGTL 30
Db 185 AGQKVGTL 192

RESULT 13
ADY37608
ID ADY37608 standard; protein; 494 AA.
XX
AC ADY37608;
XX
DT 19-MAY-2005 (first entry)
XX
DE Lung cancer related protein, SEQ ID 84.
XX
KW Cytostatic; lung tumor.
XX
OS Homo sapiens.
XX
PN US2005048589-A1.
XX
PD 03-MAR-2005.
XX
PF 25-AUG-2004; 2004US-00926543.
XX
PR 25-AUG-2003; 2003US-0497790P.
XX
PA (JEND/) JENDOUBI M.
XX
PI Jendoubi M;
XX
DR WPI; 2005-221517/23.
DR N-PSDB; ADY37562.
XX
XX New antibody specific for an antigen differentially expressed between
PT lung cancer tissue and normal tissue, useful for diagnosing, treating or
PT preventing lung cancer.
XX
PS Disclosure; SEQ ID NO 84; 171pp; English.
XX
CC The present invention relates to an antibody (I) that binds to an antigen
CC that is differentially expressed between lung cancer tissue and normal
CC tissue, where the antigen is encoded by ADY37525. Also disclosed are
CC polypeptides (ADY37573-ADY37618) encoded by lung cancer related genes
CC (ADY37525-ADY37572), their fragments, analogs or derivatives; producing
CC lung cancer related polypeptides; and a diagnostic kit for detection and
CC disease management of lung cancer. (I) is useful for characterizing and
CC analyzing biological activity and function of the lung cancer-specific
CC gene products in relation to cellular pathways and networks in normal and
CC disease states, purifying the cancer specific gene products, and
CC detecting lung cancer-specific gene products and their expression levels
CC in animal models of cancer. (I) is also useful for diagnosing, treating
CC or preventing lung cancer, and also disorders related to abnormal
CC cellular differentiation, proliferation or degeneration.
XX
SQ Sequence 494 AA;
Query Match 6.2%; Score 8; DB 9; Length 494;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 23 AGQKVGTL 30
Db 185 AGQKVGTL 192

RESULT 14
ABU62853
ID ABU62853 standard; protein; 559 AA.
XX
AC ABU62853;
XX
DT 09-SEP-2003 (first entry)
XX
DE Novel endochitinase.
XX
KW Endochitinase; GM13 strain; chitin oligosaccharide; NAG; chitin.
XX

OS Cellulomonas sp.
 XX RN2002075143-A.
 XX
 PD 04-OCT-2002.
 XX
 XX 23-MAR-2001; 2001KR-00015379.
 XX
 XX 23-MAR-2001; 2001KR-00015379.
 XX
 XX (AMIC-) AMICOGEN CO LTD.
 XX
 XX Choi YJ, Jun YJ, Jung GH, Shin YC;
 XX WPI; 2003-146028/14.
 DR N-PSDB; ACD26615.
 DR
 XX Novel endochitinase gene from cellulomonas sp. gml3 strain.
 PT
 XX Claim 2; Fig 2; 7pp; Korean.
 PS
 XX The invention describes the nucleotide sequence of a novel endochitinase
 CC gene derived from Cellulomonas sp. Gml3 strain. Also, provided is the
 CC amino acid sequence translated therefrom. Therefore, chitin
 CC oligosaccharide and NAG can be manufactured from chitin by using an
 CC endochitinase encoded by the novel endochitinase gene. This is the amino
 CC acid sequence of the novel endochitinase protein
 XX
 XX Sequence 559 AA;
 SQ
 Query Match 6.2%; Score 8; DB 6; Length 559;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 114 AGTGTATF 121
 DB 45 AGTGTATF 52
 RESULT 15
 ABG29271
 ID ABG29271 standard; protein; 574 AA.
 XX
 AC ABG29271;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #29262.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 FN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 FI
 XX WPI; 2001-639362/73.
 DR
 XX N-PSDB; AAS93458.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 59630; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 574 AA;
 SQ
 Query Match 6.2%; Score 8; DB 4; Length 574;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 AGQKVGTL 30
 DB 308 AGQKVGTL 315
 Search completed: April 26, 2006, 17:09:41
 Job time : 189 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:06:39 ; Search time 16 Seconds
(without alignments)
769.734 Million cell updates/sec

Title: US-10-018-892-3
Perfect score: 128
Sequence: 1 AAQNTTSANWSQDPGFTGPA.....KSTLPAGTFTATFYVQYQN 128
Scoring Table: ~~obscure~~
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues
Word size: 1
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database : PIR_80.*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	165	A40618	fimbrin, SEF14 - S
2	7	5.5	100	C64431	conserved hypothetical
3	7	5.5	104	JC7376	hardening-inducibl
4	7	5.5	188	H95348	nitric-oxide reduc
5	7	5.5	242	C84380	hypothetical prote
6	7	5.5	247	A55208	short-chain alcoh
7	7	5.5	248	T12632	water channel prot
8	7	5.5	248	T01648	probable tonoplast
9	7	5.5	248	T07819	hypothetical water cha
10	7	5.5	277	E82739	hypothetical prote
11	7	5.5	292	H87260	asparaginase famil
12	7	5.5	338	T32972	hypothetical prote
13	7	5.5	369	AE2345	phospho-N-acetylmu
14	7	5.5	390	F95350	VirB10 transmembra
15	7	5.5	398	AI3231	conjugal transfer
16	7	5.5	417	E95327	conserved hypothet
17	7	5.5	434	T47772	hypothetical prote
18	7	5.5	454	AI1278	anthranilate synth
19	7	5.5	454	AI1641	anthranilate synth
20	7	5.5	471	T41407	hypothetical prote
21	7	5.5	497	A29055	aldehyde dehydroge
22	7	5.5	512	H86206	hypothetical prote
23	7	5.5	513	T45456	NADH2 dehydrogenas
24	7	5.5	515	E84577	probable cycokinin
25	7	5.5	526	T70539	hypothetical prote
26	7	5.5	534	T84398	phosphate ABC tran
27	7	5.5	543	T32109	hypothetical prote
28	7	5.5	580	S56645	chaperonin 60 alph
29	7	5.5	594	Q96525	protein Tln15.25 [

30 7 5.5 594 2 T33686 hypothetical prote
31 7 5.5 715 2 S76492 lipoprotein nlpd -
32 7 5.5 721 2 D83110 exodeoxyribonuclea
33 7 5.5 802 2 AF1227 phenylalanyl-tRNA
34 7 5.5 802 2 AH1580 phenylalanyl-tRNA
35 7 5.5 808 2 F84038 phenylalanyl-tRNA
36 7 5.5 818 2 T33689 hypothetical prote
37 7 5.5 818 2 T33687 hypothetical prote
38 7 5.5 826 1 D64978 probable outer mem
39 7 5.5 826 2 C90993 probable outer mem
40 7 5.5 826 2 F85838 probable outer mem
41 7 5.5 863 2 S37088 phycobilisome anch
42 7 5.5 1018 2 H83135 probable adhesin P
43 7 5.5 1167 2 B75258 probable S-layer p
44 7 5.5 1336 2 T17479 hypothetical prote
45 7 5.5 1337 1 I38670 protein-tyrosine-p

ALIGNMENTS

RESULT 1
A40618
fimbrin, SEF14 - Salmonella enteritidis
N:Alternate names: fimbrin, Sefa
C:Species: Salmonella enteritidis
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A40618, A25034
R:Clouthier, S.C.; Muller, K.H.; Doran, J.L.; Collinson, S.K.; Kay, W.W.
J. Bacteriol. 175, 2523-2533, 1993
A:Title: Characterization of three fimbrial genes, sefABC, of Salmonella enteritidis.
A:Reference number: A40618; MUID:93239677; PMID:8097515
A:Contents: 27655-3b
A:Accession: A40618
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-165 <CLO>
A:Cross-references: UNIPROT:PI2061; UNIPARC:UPI00001135BD; GB:L11008; NID:g310645; PID
A:Note: sequence extracted from NCBI backbone (NCBIN:130387, NCBIP:130395)
R:Feutrier, J.; Kay, W.W.; frust, T.J.
J. Bacteriol. 168, 221-227, 1986
A:Title: Purification and characterization of fimbriae from Salmonella enteritidis.
A:Reference number: A25034; MUID:87008384; PMID:2875990
A:Accession: A25034
A:Molecule type: protein
A:Residues: 22-78, 'X', 80-82, 'XQ', 128 <FEU>
A:Cross-references: UNIPARC:UPI00001798D8

Query Match 100.0%; Score 128; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-119; Indels 0; Gaps 0;
Matches 128; Conservative 0; Mismatches 0;

QY 1 AAQNTTSANWSQDPGFTGPAVAAQKVGKGLSLTATGPHNSVSIAGKASVSGGVATVPFV 60
Db 38 AAQNTTSANWSQDPGFTGPAVAAQKVGKGLSLTATGPHNSVSIAGKASVSGGVATVPFV 97
QY 61 DQGGQPVFRGRIQGANINDQANTGIDGLAGRWASQETLNVPTTFGKSLTLPAGTPTAT 120
Db 98 DQGGQPVFRGRIQGANINDQANTGIDGLAGRWASQETLNVPTTFGKSLTLPAGTPTAT 157

QY 121 FVVOQYQN 128
Db 158 FVVOQYQN 165

RESULT 2
C64431
conserved hypothetical protein MJ1052 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: C64431
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8698087
A:Accession: C64431
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-100 <BUL>
A:Cross-references: UNIPROT:Q58452; UNIPARC:UPI0000139F75; GB:U67548; GB:L77117; NID:928
C:Genetics:
A:Map position: REV989804-989502
A:Start codon: GNG
C:Superfamily: *Methanococcus jannaschii* conserved hypothetical protein MJ1052

Query Match 5.5%; Score 7; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GKGSVS 51
| | | | |
DB 17 GKGSVS 23

RESULT 3
JC7376
hardening-inducible protein HIC12 - *Chlorella vulgaris*
C:Species: *Chlorella vulgaris*
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C:Accession: JC7376
R:Honjoh, K.; Matsumoto, H.; Shimizu, H.; Ooyama, K.; Tanaka, K.; Oda, Y.; Takata, R.; J
Biochem. Biotechnol. Biochem. 64, 1656-1663, 2000
A:Title: Cryptoreactive activities of group 3 late embryogenesis abundant proteins from
A:Reference number: JC7376
A:Accession: JC7376
A:Molecule type: mRNA
A:Residues: 1-104 <HON>
A:Cross-references: UNIPROT:Q9SLP7; UNIPARC:UPI00000A7366; DDBJ:AB035642
A:Experimental source: strain IAM C-27
C:Comment: This protein, homolog of group 3 late embryogenesis abundant protein, has eff
C:Genetics:
A:Gene: hic12

Query Match 5.5%; Score 7; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AAGQKVG 28
| | | | |
DB 17 AAGQKVG 23

RESULT 4
H95348
nitric-oxide reductase (EC 1.7.99.7) NorE [imported] - *Sinorhizobium meliloti* (strain 10
C:Species: *Sinorhizobium meliloti*
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: H95348
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9886, 2001
A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: H95348
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <KUR>
A:Cross-references: UNIPROT:Q92214; UNIPARC:UPI00000CB18E; GB:AE006469; PIDN:AAK65354.1;
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hymen, R.W.; Jones, T.
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
heault, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: norE
A:Genome: plasmid
C:Keywords: oxidoreductase

Query Match 5.5%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AGKGASV 50
| | | | |
DB 84 AGKGASV 90

RESULT 5
C84380
hypothetical protein hisA [imported] - *Halobacterium* sp. NRC-1
C:Species: *Halobacterium* sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84380
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabb
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; I
A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84380
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <STO>
A:Cross-references: UNIPROT:Q9HN14; UNIPARC:UPI000012C7C5; GB:AE004437; NID:gl0581706;
C:Genetics:
A:Gene: hisA

Query Match 5.5%; Score 7; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 SGGVATV 57
| | | | |
DB 200 SGGVATV 206

RESULT 6
A55208
short-chain alcohol dehydrogenase homolog socA1 - *Myxococcus xanthus*
C:Species: *Myxococcus xanthus*
C:Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 05-Oct-2004
C:Accession: A55208
R:Ree, K.; Shinkets, L.J.
J. Bacteriol. 176, 2200-2209, 1994
A:Title: Cloning and characterization of the socA locus which restores development to
A:Reference number: A55208; MUID:94209218; PMID:8157590
A:Accession: A55208
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <LEB>
A:Cross-references: UNIPROT:Q7M0Y6; UNIPARC:UPI0000174E88; GB:L27430
C:Genetics:
A:Gene: socA1
C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: alcohol metabolism
F:7-187/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 5.5%; Score 7; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 TGPAAVAA 23
|||||
Db 197 TGPAAVAA 203

RESULT 7

T12632
water channel protein - common sunflower
N:Alternate names: aquaporin; tonoplast intrinsic protein homolog
C:Species: Helianthus annuus (common sunflower)
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 05-Oct-2004
C:Accession: T12632
R:Sarda, X.; Touch, D.; Ferrare, K.; Legrand, E.; Dupuis, J.M.; Casse-Delbart, P.; Lama
Plant J. 12, 1103-1111, 1997
A>Title: Two TIP-like genes encoding aquaporins are expressed in sunflower guard cells.
A:Reference number: Z17561; MUID:98079246; PMID:9418051
A:Accession: T12632
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-248 <SAR>
A:Cross-references: UNIPROT:Q39935; UNIPARC:UPI00000A08DF; EMBL:X95952; NID:g1212914; PI
A:Experimental source: root
C:Superfamily: aquaporin
C:Keywords: channel-forming protein; transmembrane protein

Query Match 5.5%; Score 7; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GPAVAAAG 24
|||||
Db 203 GPAVAAAG 209

RESULT 8

T01648
probable tonoplast aquaporin - maize
C:Species: Zea mays (maize)
C>Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 05-Oct-2004
R:Finkelstein, D.B.
C:Accession: T01648
submitted to the EMBL Data Library, April 1998
A:Reference number: Z14367
A:Accession: T01648
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-248 <PIN>
A:Cross-references: UNIPROT:O81216; UNIPARC:UPI00000ACCBF; EMBL:AF057183; NID:g3264595;
A:Experimental source: strain TX5855
C:Superfamily: aquaporin

Query Match 5.5%; Score 7; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GPAVAAAG 24
|||||
Db 202 GPAVAAAG 208

RESULT 9

T07819
probable water channel protein delta-VM23 - radish
C:Species: Raphanus sativus (radish)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 05-Oct-2004
C:Accession: T07819
R:Higuchi, T.; Suga, S.; Teuchiya, T.; Hisada, H.; Morishima, S.; Okada, Y.; Maeshima, M
Plant Cell Physiol. 39, 905-913, 1998
A>Title: Molecular cloning, water channel activity and tissue specific expression of two
A:Reference number: Z16148; MUID:99033463; PMID:9816675
A:Accession: T07819
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-248 <HIG>
A:Cross-references: UNIPROT:O82142; UNIPARC:UPI00000A5347; EMBL:AB010416; NID:g3298326
A:Experimental source: cultivar Tokinashidaikon
C:Genetics:
A:Gene: VIP3
C:Superfamily: aquaporin
C:Keywords: water channel

Query Match 5.5%; Score 7; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GPAVAAAG 24
|||||
Db 203 GPAVAAAG 209

RESULT 10

E82739
hypothetical protein XF0968 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ
C:Accession: E82739
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: E82739
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-277 <SIM>
A:Cross-references: UNIPROT:Q9PER0; UNIPARC:UPI00000C25A7; GB:AE003935; GB:AE003849; N
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fro
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Lai
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A:Authors: Martins, E.M.P.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, B.C.; Miyaki, C.V.
P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawas
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tshahako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0968

Query Match 5.5%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AQTNTSA 8
|||||
Db 83 AQTNTSA 89

RESULT 11

H87260
asparaginase family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87260
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, M
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.H.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: H87260
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-292 <STO>
A;Cross-references: UNIPROT:Q9ABX3; UNIPARC:UPI000000C6F20; GB:AE005673; NID:gl3421200; F
C;Genetics:
A;Gene: CC0097
C;Superfamily: ybiK protein

Query Match 5.5%; Score 7; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 IAGKGAS 49
|||||
Db 64 IAGKGAS 70

RESULT 12
T32972
hypothetical protein F58H7.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32972
R;Graves, T.
submitted to the EMBL Data Library, February 1998
A;Description: The sequence of C. elegans cosmid F58H7.
A;Reference number: T32972
A;Accession: T32972
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-338 <GRA>
A;Cross-references: UNIPROT:O45094; UNIPARC:UPI000007F204; EMBL:AF045643; PIDN:AAC02595.
A;Experimental source: strain Bristol N2; clone F58H7
C;Genetics:
A;Gene: CESP:F58H7.8
A;Map position: 4
A;Introns: 31/2
C;Superfamily: Caenorhabditis elegans hypothetical protein F42G2.3

Query Match 5.5%; Score 7; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 KVGTLIS 32
|||||
Db 183 KVGTLIS 189

RESULT 13
AE2345
phospho-N-acetylmuramoyl-pentapeptide-transferase e [imported] - Nostoc sp. (strain PCC 7
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE2345
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2345
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-369 <KUR>
A;Cross-references: UNIPROT:Q8YP83; UNIPARC:UPI000012F50C; GB:BA000019; PIDN:BAE76015-1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all4316
C;Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase

Query Match 5.5%; Score 7; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 50;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GIDGLAG 90
|||||
Db 214 GIDGLAG 220

RESULT 14
F95350
VirB10 transmembrane type IV secretion protein [imported] - Sinorhizobium meliloti (str
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95350
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
; Kalman, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melil
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: F95350
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-390 <KUR>
A;Cross-references: UNIPROT:Q92202; UNIPARC:UPI00000CB19A; GB:AE006469; PIDN:AAK65368.
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lela
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A36039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: virB10
A;Genome: plasmid
C;Superfamily: tumor-inducing plasmid pTic58 virB10 protein

Query Match 5.5%; Score 7; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AQTNTSA 8
|||||
Db 314 AQTNTSA 320

RESULT 15
AI3231
conjugal transfer protein trbL [imported] - Agrobacterium tumefaciens (strain C58, Dup
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AI3231
R;Wood, D.W.; Stetubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AI3231
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-398 <KUR>
A;Cross-references: UNIPROT:O61175; UNIPARC:UPI00000D1462; GB:AE008690; PIDN:AAL46271.
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: trbL
A;Genome: plasmid

Query Match 5.5%; Score 7; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 53;

	Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	47	GASVSGG 53								
Db	282	GASVSGG 288								

Search completed: April 26, 2006, 17:10:01
Job time : 18 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:06:38 ; Search time 163 Seconds
(without alignments)

554.034 Million cell updates/sec

Title: US-10-018-892-3

Perfect score: 128

Sequence: 1 AAQNTTSANWSQDPGTGPA.....KSTLPAGTFTATFYVQYQVN 128

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2166443 seqs, 705528306 residues

Word-size: 1

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	165	1 FM SALEN	P12061 salmonella
2	73	57.0	165	2 Q5PM43 SALPA	Q5pm43 salmonella
3	10	7.8	30	2 Q9R4V0 SALEN	Q9r4v0 salmonella
4	9	7.0	606	2 Q82KE7 STRAW	Q82ke7 streptomyc
5	9	7.0	725	2 Q8XZRI RALSO	Q8xzi1 ralistonia s
6	9	7.0	741	2 Q9X6X9 BURPI	Q9x6x9 burkholderi
7	8	6.2	187	2 Q4PLV2 USTMA	Q4ply2 ustilago ma
8	8	6.2	205	2 Q73TEL MYCPA	Q73tel mycobacteri
9	8	6.2	310	2 Q6ZLKI ORYSA	Q6zlk1 oryza sativ
10	8	6.2	329	2 Q88SS6 LACPL	Q88ss6 lactobacill
11	8	6.2	355	2 Q9ZBD2 TREDE	Q9zbd2 treponema d
12	8	6.2	411	2 Q95LJ3 MACFA	Q95lj3 macaca fasc
13	8	6.2	453	2 Q8DVF8 STRMU	Q8dvf8 streptococ
14	8	6.2	456	2 Q5LYI2 STRT1	Q5lyi2 streptococ
15	8	6.2	456	2 Q5M345 STRT2	Q5m345 streptococ
16	8	6.2	512	2 Q4HZE9 GIBZE	Q4hze9 gibberella
17	8	6.2	528	2 Q6LXK8 PHOPR	Q6lrx8 photobacter
18	8	6.2	558	2 Q4R6U5 MACFA	Q4r6u5 macaca fasc
19	8	6.2	559	2 Q4BN2_9CELL	Q4bn2 cellulomona
20	8	6.2	579	1 ESR1 SPAAU	Q9pvz9 sparus aura
21	8	6.2	583	2 Q804Q6 ACASC	Q804q6 acanthopagr
22	8	6.2	587	2 Q74KW2 LACJO	Q74kw2 lactobacill
23	8	6.2	599	1 RSNL2 RAT	Q66hd5 rattus norv
24	8	6.2	614	2 Q4ZFY2 HUMAN	Q4zfy2 homo sapien
25	8	6.2	639	2 Q8VP60 TREDE	Q8vp60 treponema d
26	8	6.2	704	1 RSNL2 MOUSE	Q8ci96 mus musculu
27	8	6.2	705	1 RSNL2 HUMAN	Q8n3c7 homo sapien
28	8	6.2	716	2 Q4SK06 TETNG	Q4sk06 tetraodon n
29	8	6.2	724	2 Q59GR7 HUMAN	Q59gr7 homo sapien
30	7	5.5	26	2 Q9QRU9_9HEPC	Q9qru9 hepatitis c
31	7	5.5	26	2 Q9QRV0_9HEPC	Q9qrv0 hepatitis c

```

32 7 5.5 66 2 Q6Z7W2 ORYSA
33 7 5.5 69 2 Q67TN3 ORYSA
34 7 5.5 86 2 Q6K202 ORYSA
35 7 5.5 94 2 Q6LD31 LYMDI
36 7 5.5 100 1 Y1052 METJA
37 7 5.5 104 2 Q9SLP7 CHLVU
38 7 5.5 106 2 P95679 COXBU
39 7 5.5 106 2 Q8GM40 LEGFN
40 7 5.5 114 1 CT149 HUMAN
41 7 5.5 115 1 CT149 MOUSE
42 7 5.5 115 2 Q5PRO1 RAT
43 7 5.5 115 2 Q58EU5 MOUSE
44 7 5.5 120 2 Q6DIQ8 ERWCT
45 7 5.5 121 1 CHA9 LYMDI

```

ALIGNMENTS

```

RESULT 1
FM SALEN
ID FM SALEN STANDARD; PRT; 165 AA.
AC P12061;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Fimbrial protein precursor.
GN Name=sefa; Synonyms=sef14;
OS Salmonella enteritidis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=592;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=27655-3B;
RX MEDLINE=93239677; PubMed=8097515;
RA Clouthier S.C., Mueller K.-H., Doran J.L., Collinson S.K., Kay W.W.;
RT "Characterization of three fimbrial genes, sefABC, of Salmonella
   enteritidis."
RL J. Bacteriol. 175:2523-2533 (1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=91072589; PubMed=1701443;
RA Thorns C.J., Sojka M.G., Chasey D.C.;
RT "Detection of a novel fimbrial structure on the surface of Salmonella
   enteritidis by using a monoclonal antibody."
RL J. Clin. Microbiol. 28:2409-2414 (1990).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Ogumiyi A.D., Kotlarski I., Morona R., Manning P.A.;
RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PROTEIN SEQUENCE OF 22-85.
RX MEDLINE=87008384; PubMed=2875990;
RA Feutrier J., Kay W.W., Trust T.J.;
RT "Purification and characterization of fimbriae from Salmonella
   enteritidis."
RL J. Bacteriol. 168:221-227 (1986).
CC -!- FUNCTION: Structural subunit of the sef14 fimbriae (S.enteritidis
   filamentous fimbriae).
CC -!- SUBCELLULAR LOCATION: Fimbria.

```

```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use as long as its content is in no way modified and this statement is not
   removed.
CC
CC EMBL; L11008; AAA27219.1; -; Genomic DNA.
CC EMBL; L03833; AAA71892.1; -; Unassigned DNA.
CC EMBL; X98516; CAA67141.1; -; Genomic DNA.
CC PIR; A40618; A40618.
CC PDB; 1LU0; Model; A=1-165.

```

```
DR InterPro; IPR010498; SEF14 adhesin.
DR Pfam; PF06443; SEF14 adhesin; 1.
KW 3D-structure; Direct protein sequencing; Fimbria; Signal.
FT SIGNAL 1 21
FT CHAIN 22 165 Fimbrial protein.
FT CONFLICT 30 30 V -> E (in Ref. 2 and 3).
FT CONFLICT 84 85 GA -> QW (in Ref. 4).
FT SEQUENCE 165 AA; 16477 MW; 5B33798A3F0F9091 CRC64;
SQ
Query Match 100.0%; Score 128; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.4e-119;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANWSQPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
Db 38 AAQNTTSANWSQPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 97

Qy 61 DGGQVPFRGRIGRIQANINDQANTGIDGLAGRWASSQETLNVPVTTFGKSTLPAGTTAT 120
Db 98 DGGQVPFRGRIGRIQANINDQANTGIDGLAGRWASSQETLNVPVTTFGKSTLPAGTTAT 157

Qy 121 FYVQYQYN 128
Db 158 FYVQYQYN 165

RESULT 2
QSPM43 SALPA
ID QSPM43 SALPA PRELIMINARY; PRT; 165 AA.
AC QSPM43;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Fimbrial structural protein.
GN Name=sefA; OrderedLocNames=SPA4304;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaanty K., Fronick C., Megrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP0000026; ANV80033.1; -; Genomic_DNA.
DR InterPro; IPR010498; SEF14 adhesin.
DR Pfam; PF06443; SEF14 adhesin; 1.
DR Complete proteome.
SQ SEQUENCE 165 AA; 16665 MW; 8A32BE3F43C91520 CRC64;

Query Match 57.0%; Score 73; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.1e-64;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANWSQDPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 60
Db 38 AAQNTTSANWSQDPGFTGPAVAAGKVGTLISITATGPHNSVSIAGKASVSGGVATVPFV 97

Qy 61 DGGQVPFRGRIGRIQ 73
Db 98 DGGQVPFRGRIGRIQ 110

Query Match 7.0%; Score 9; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 SVSGGVATV 57

Q9R4V0 SALEN
ID Q9R4V0 SALEN PRELIMINARY; PRT; 30 AA.
AC Q9R4V0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Fimbrial protein SEF14 (Pragmat).
OS Salmonella enteritidis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=592;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=95048770; PubMed=7960117;
RA Ogunniyi A.D., Manning P.A., Kotlarski I.;
RT "A Salmonella enteritidis 11kX pilin induces strong T-lymphocyte
RT responses.";
RL Infect. Immun. 62:5376-5383(1994).
DR InterPro; IPR010498; SEF14 adhesin.
DR Pfam; PF06443; SEF14 adhesin; 1.
DR SEQUENCE 30 AA; 3074 MW; AE75376BE8860C34 CRC64;

Query Match 7.8%; Score 10; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAQNTTSANW 10
Db 17 AAQNTTSANW 26

RESULT 4
Q82KE7 STRAW
ID Q82KE7 STRAW PRELIMINARY; PRT; 606 AA.
AC Q82KE7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=SAV2456;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.21143198;
RA Shinoe S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC70167.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 606 AA; 64012 MW; F6DBC9DE2B621818 CRC64;

Query Match 7.0%; Score 9; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 SVSGGVATV 57
```

```

DR GO: 0016787; F:hydrolase activity; IEA.
SQ SEQUENCE 741 AA; 75590 MW; 68E13C032382B373 CRC64;

Query Match 7.0%; Score 9; DB 2; Length 741;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 SVSGGVATV 57
DB 729 SVSGGVATV 737

RESULT 7
Q4PLY2 USTMA
ID Q4PLY2_USTMA PRELIMINARY; PRT; 187 AA.
AC Q4PLY2;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=UM05881.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nussbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Bowersky M., Boukngalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseles M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulkarni B., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihailev A., Mihova T., Mikkelsen T., Mienga V., Moru K.,
RA Moses J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okaowo O., O'leary S., Omotosho B.,
RA O'Neill K., Oman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sounez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuanga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Tovey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataravan V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee B., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Zander E.;
RT "The genome sequence of Ustilago maydis."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC CC
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACF01000215; EAK86826.1; -; Genomic_DNA.

```

KW Hypothetical protein.
SQ SEQUENCE 187 AA; 20885 MW; AF6B8CB7C85D14E2 CRC64;

Query Match 6.2%; Score 8; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 VSIAGKGA 48
| | | | | | | |
Db 40 VSIAGKGA 47

RESULT 8
Q73TEL_MYCPA
ID Q73TEL_MYCPA PRELIMINARY; PRT; 205 AA.
AC Q73TEL;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=MAP3777;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L.; Banantaine J.; Zhang Q.; Amonsin A.; Alt D.; Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017240; AAS06327.1; -; Genomic_DNA.
DR InterPro; IPR011610; Mtu_fam_121.
DR InterPro; IPR003455; Ont_N.
DR Pfam; PF02409; Ont_N; 1.
DR TIGRFAMs; TIGR00027; mthyl_TIGR00027; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 205 AA; 22413 MW; 80A09C60B6237F10 CRC64;

Query Match 6.2%; Score 8; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 TGIDGLAG 90
| | | | | | | |
Db 100 TGIDGLAG 107

RESULT 9
Q6ZLK1_ORYSA
ID Q6ZLK1_ORYSA PRELIMINARY; PRT; 310 AA.
AC Q6ZLK1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Membrane transporter PFB0275w-like protein.
GN Name=QJ1136_A05.20;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
clone:QJ1136_A05.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003748; BAC82966.1; -; Genomic_DNA.
DR Gramene; O6ZLK1; -;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011701; MFS_1.

DR Pfam; PF07690; MFS_1; 1.
SQ SEQUENCE 310 AA; 34286 MW; B69FC90CFFAFD5EC CRC64;

Query Match 6.2%; Score 8; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 GLAGWRVA 94
| | | | | | | |
Db 8 GLAGWRVA 15

RESULT 10
Q8SS6_LACPL
ID Q8SS6_LACPL PRELIMINARY; PRT; 329 AA.
AC Q8SS6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Esterase (Putative).
GN OrderedLocNames=lp_3312;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCIMB 8826 / WCFS1;
RA MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
RA Kleerebezem M.; Boekhorst J.; van Kranenburg R.; Molenaar D.;
RA Kuipers O.P.; Leer R.; Tarchini R.; Peters S.A.; Sandbrink H.M.;
RA Fiers M.W.E.J.; Stiekema W.; Klein Lankhorst R.M.; Bron P.A.;
RA Hoffer S.M.; Nierop Groot M.N.; Kerkhoven R.; De Vries M.; Ursing B.;
RA De Vos W.M.; Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR EMBL; AL935261; CAD65434.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0046677; P:response to antibiotic; IEA.
DR InterPro; IPR001466; Beta_lactamase.
DR Pfam; PF00144; Beta_lactamase; 1.
KW Complete proteome.
SQ SEQUENCE 329 AA; 37332 MW; FC66021317FCC1B7 CRC64;

Query Match 6.2%; Score 8; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GFTGPAVA 22
| | | | | | | |
Db 283 GFTGPAVA 290

RESULT 11
Q9ZBD2_TREDE
ID Q9ZBD2_TREDE PRELIMINARY; PRT; 355 AA.
AC Q9ZBD2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 43kDa protein.
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC35405;
RX MEDLINE=97101036; PubMed=8945563;
RA Ishihara K.; Miura T.; Kuramitsu H.K.; Oduda K.;
RT "Characterization of the Treponema denticola prp gene encoding a
prolyl-phenylalanine-specific protease (denticillin).";
RL Infect. Immun. 64:5178-5186(1996).
RN [2]

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC35405;
RA Iehihara K., Miura T., Kuramitsu H.K., Okuda K.;
RT "Characterization of the Treponema denticola prtp gene expressing a
RL sublyti-phenylalanine specific protease (dentilisins).";
DR EMBL; D83264; BAAL1873.1; -; Genomic DNA.
SQ SEQUENCE 355 AA; 38855 MW; 29DD36C1B9551A8 CRC64;

Query Match          6.2%; Score 8; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 DGGQGPVF 68
DB 188 DGGQGPVF 195

RESULT 12
Q95LJ3_MACFA PRELIMINARY; PRT; 411 AA.
AC Q95LJ3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Osada N., Hida M., Kusuda J., Tanuma R., Hirata M., Suto Y., Hirai M.,
RA Terao K., Sugano S., Hashimoto K.;
RT "Cynomolgus monkey testicular cDNAs for discovery of novel human genes
RT in the human genome sequence.";
RL BMC Genomics 3:36-36(2002).
DR EMBL; AB072795; BAB69764.1; -; mRNA.
DR HSSP; Q20728; 1LPL.
DR SMR; Q95LJ3; 159-247.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000938; CAP-Gly.
DR Pfam; PF00023; ANK; 3.
DR Pfam; PF01302; CAP_GLY; 2.
DR SMART; SM00248; ANK; 2.
DR PROSITE; PS0297; ANK_REPEAT; 1.
DR PROSITE; PS0088; ANK_REPEAT; 1.
DR PROSITE; PS00845; CAP_GLY_1; 1.
DR PROSITE; PS0245; CAP_GLY_2; 2.
DR ANK repeat; Hypothetical protein; Repeat.
KW ANK repeat; Hypothetical protein; Repeat.
SQ SEQUENCE 411 AA; 44660 MW; AA0D8543701FD11B CRC64;

Query Match          6.2%; Score 8; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AGQKVGTL 30
DB 185 AGQKVGTL 192

RESULT 13
Q8DVFB_STRMU PRELIMINARY; PRT; 453 AA.
AC Q8DVFB;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative anthranilate synthase, alpha subunit (EC 4.1.3.27).
GN Name=trpE; OrderedLocNames=SMU.532;
OS Streptococcus mutans.

```

```

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=UAI59 / ATCC 700610 / Serotype c;
RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
RA Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014898; AAN58275.1; -; Genomic DNA.
DR HSSP; P05041; 1KOG.
DR GO; GO:0004049; P:anthranilate synthase activity; IEA.
DR GO; GO:0009058; P:biogenesis; IEA.
DR GO; GO:0000162; P:tryptophan biosynthesis; IEA.
DR InterPro; IPR005256; Anth_synth.
DR InterPro; IPR005801; Anth_synth_chor.
DR InterPro; IPR006805; Anth_synth_I_N.
DR Pfam; PF04715; Anth_synth_I_N; 1.
DR Pfam; PF00425; Chorismate_Bind; 1.
DR PRINTS; PR00095; ANTSNTHASEI.
DR ProDom; PD000779; Anth_synth_chor; 1.
DR TIGRFAMs; TIGR00564; trpE_m0st; 1.
KW Complete proteome.
SQ SEQUENCE 453 AA; 51042 MW; DECD467AB4A8A0D CRC64;

Query Match          6.2%; Score 8; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 KSTLPAGT 116
DB 358 KSTLPAGT 365

RESULT 14
Q5LYI2_STRT1 PRELIMINARY; PRT; 456 AA.
AC Q5LYI2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Anthranilate synthase component I.
GN Name=trpE; OrderedLocNames=stl1593;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquin B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goltsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR EMBL; CP000024; AAV63123.1; -; Genomic DNA.
DR GO; GO:0004049; P:anthranilate synthase activity; IEA.
DR GO; GO:0009058; P:biogenesis; IEA.
DR GO; GO:0000162; P:tryptophan biosynthesis; IEA.
DR InterPro; IPR005256; Anth_synth.
DR InterPro; IPR005801; Anth_synth_chor.
DR InterPro; IPR006805; Anth_synth_I_N.
DR Pfam; PF04715; Anth_synth_I_N; 1.
DR Pfam; PF00425; Chorismate_Bind; 1.
DR PRINTS; PR00095; ANTSNTHASEI.

```

```

DR ProDom; PD000779; Anth_synth_chor; 1.
DR TIGRFAMs; TIGR00564; trpE_most; 1.
KW Complete proteome.
SQ SEQUENCE 456 AA; 51170 MW; 5968B272A684A1A6 CRC64;

Query Match      6.2%; Score 8; DB 2; Length 456;
Best Local Similarity 100.0%; Pred.No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 KSTLPAGT 116
Db 361 KSTLPAGT 368

RESULT 15
Q5M345_STRT2
ID Q5M345_STRT2 PRELIMINARY; PRT; 456 AA.
AC Q5M345;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Anthranilate synthase component I.
GN Name=trpE; OrderedLocNames=stul593;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goltsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtiau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558 (2004).
DR EMBL; CP000023; AAV61196.1; -; Genomic DNA.
DR GO; GO:0004049; P:anthranilate synthase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0000162; P:tryptophan biosynthesis; IEA.
DR InterPro; IPR005256; Anth_synthI.
DR InterPro; IPR005801; Anth_synth_chor.
DR InterPro; IPR006805; Anth_synth_I_N.
DR Pfam; PF04715; Anth_synt_I_N; 1.
DR Pfam; PF00425; Chorismate_bind; 1.
DR PRINTS; PR00095; ANTSYNTHASEI.
DR ProDom; PD000779; Anth_synth_chor; 1.
DR TIGRFAMs; TIGR00564; trpE_most; 1.
KW Complete proteome.
SQ SEQUENCE 456 AA; 51170 MW; 5968B272A684A1A6 CRC64;

```

```

Query Match      6.2%; Score 8; DB 2; Length 456;
Best Local Similarity 100.0%; Pred.No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 109 KSTLPAGT 116
Db 361 KSTLPAGT 368

```

```

Search completed: April 26, 2006, 17:09:49
Job time : 164 secs

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:07:08 ; Search time 47 Seconds
(without alignments)
225.159 Million cell updates/sec

Title: US-10-018-892-3
Perfect score: 128
Sequence: 1 AAQNTTSANWSQDPGFTGPA.....KSTLPAGTFTATFYVQYQN 128

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size : 1

Total number of hits satisfying chosen parameters: 566820

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgm2_6/ptodata/1/iaa/5 COMB.pep:*
 - 2: /cgm2_6/ptodata/1/iaa/6 COMB.pep:*
 - 3: /cgm2_6/ptodata/1/iaa/H_COMB.pep:*
 - 4: /cgm2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
 - 5: /cgm2_6/ptodata/1/iaa/RE_COMB.pep:*
 - 6: /cgm2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	144	2	US-09-230-078A-6
2	128	100.0	165	1	US-08-233-788A-40
3	128	100.0	176	1	US-08-449-922-1
4	128	100.0	180	2	US-09-230-078A-4
5	115	89.8	143	2	US-09-543-407-46
6	7	5.5	17	2	US-09-688-017-194
7	7	5.5	17	2	US-09-688-017-217
8	7	5.5	114	2	US-09-443-184-53
9	7	5.5	132	2	US-09-125-642C-15
10	7	5.5	132	2	US-09-431-888-11
11	7	5.5	143	1	US-08-470-179-28
12	7	5.5	239	2	US-09-372-422A-42
13	7	5.5	245	2	US-09-902-540-13474
14	7	5.5	247	2	US-09-372-422A-48
15	7	5.5	249	2	US-09-372-422A-22
16	7	5.5	256	2	US-09-902-540-11830
17	7	5.5	376	2	US-09-056-556-20
18	7	5.5	376	2	US-09-072-596-197
19	7	5.5	376	2	US-09-072-967-202
20	7	5.5	376	2	US-10-193-002-197
21	7	5.5	376	2	US-10-084-843-202
22	7	5.5	452	2	US-09-252-991A-25818
23	7	5.5	611	2	US-09-252-991A-32402
24	7	5.5	739	2	US-09-413-814-86
25	7	5.5	757	2	US-09-413-814-84
26	7	5.5	791	2	US-09-252-991A-23201
27	7	5.5	826	2	US-09-492-705A-286

28	7	5.5	1034	2	US-09-252-991A-26658	Sequence 26658, A
29	7	5.5	1195	2	US-09-902-540-11967	Sequence 11967, A
30	7	5.5	1337	2	US-08-854-585-2	Sequence 2, Appli
31	7	5.5	1337	2	US-09-447-533-2	Sequence 2, Appli
32	7	5.5	1337	4	PCT-US95-05512-2	Sequence 2, Appli
33	6	4.7	21	2	US-09-010-317-20	Sequence 20, Appl
34	6	4.7	43	2	US-09-057-363C-62	Sequence 62, Appl
35	6	4.7	43	2	US-09-265-107-62	Sequence 62, Appl
36	6	4.7	54	2	US-09-621-976-5718	Sequence 5718, Ap
37	6	4.7	66	2	US-09-107-433-2721	Sequence 2721, Ap
38	6	4.7	67	2	US-09-270-767-58993	Sequence 58993, A
39	6	4.7	73	2	US-09-270-767-40382	Sequence 40382, A
40	6	4.7	73	2	US-09-270-767-55598	Sequence 55598, A
41	6	4.7	77	2	US-09-252-991A-19887	Sequence 19887, A
42	6	4.7	85	2	US-09-484-577A-67	Sequence 67, Appl
43	6	4.7	88	2	US-09-107-532A-4811	Sequence 4811, Ap
44	6	4.7	92	2	US-08-905-223-408	Sequence 408, App
45	6	4.7	95	2	US-09-540-236-2993	Sequence 2993, Ap

ALIGNMENTS

RESULT 1
US-09-230-078A-6
; Sequence 6, Application US/09230078A
; Patent No. 6495334
; GENERAL INFORMATION:
; APPLICANT: Rajashekara, Gireesh
; APPLICANT: Kakambi, Nagarata V.
; APPLICANT: Kapur, Vivek
; TITLE OF INVENTION: RECOMBINANT SEF14 FIMBRIAL PROTEIN FROM SALMONELLA
; FILE REFERENCE: 600.335USWO
; CURRENT APPLICATION NUMBER: US/09/230, 078A
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: PCT/US97/12639
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: 60/022,191
; PRIOR FILING DATE: 1996-07-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Salmonella enteritidis
US-09-230-078A-6

Query Match	100.0%	Score 128;	DB 2;	Length 144;
Best Local Similarity	100.0%	Pred. No. 4.9e-115;	Mismatches 0;	Indels 0;
Matches 128;	Conservative 0;			Gaps 0;
QY	1	AAQNTTSANWSQDPGFTGPAVAAQKVGTL	1	AAQNTTSANWSQDPGFTGPAVAAQKVGTL
Db	17	AAQNTTSANWSQDPGFTGPAVAAQKVGTL	17	AAQNTTSANWSQDPGFTGPAVAAQKVGTL
QY	61	DQCGQPVFRGRITQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTPTAT	120	
Db	77	DQCGQPVFRGRITQGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTPTAT	136	
QY	121	FVYQYQYQN	128	
Db	137	FVYQYQYQN	144	

RESULT 2
US-08-233-788A-40
; Sequence 40, Application US/08233788A
; Patent No. 5635617
; GENERAL INFORMATION:
; APPLICANT: Doran, James L.
; APPLICANT: Kay, William W.
; APPLICANT: Collinson, Karen S.
; APPLICANT: Clouthier, Sharon C.

```

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTION
; TITLE OF INVENTION: OF SALMONELLA
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/233,788A
; FILING DATE: 26-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KING, Joshua
; REGISTRATION NUMBER: 35,570
; REFERENCE/DOCKET NUMBER: 920043.403C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANBERRY
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-233-788A-40

Query Match 100.0%; Score 128; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.5e-115;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAQNTTSANWSDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGASVSGVATVPFV 60
DB 38 AAQNTTSANWSDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGASVSGVATVPFV 97

QY 61 DQGGQPVFRGRIQGANINDQANTGIDGLAGNRVASSQETLNVPTTFGKSTLPAGTFTAT 120
DB 98 DQGGQPVFRGRIQGANINDQANTGIDGLAGNRVASSQETLNVPTTFGKSTLPAGTFTAT 157

QY 121 FYVQQYQN 128
DB 158 FYVQQYQN 165

RESULT 3
US-08-449-922-1
; Sequence 1, Application US/08449922
; Patent No. 5510241
; GENERAL INFORMATION:
; APPLICANT: THORNS, CHRISTOPHER J
; TITLE OF INVENTION: METHOD OF TESTING FOR SALMONELLA
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P C
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/449,922
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08,030,208
; FILING DATE: 26-MAR-1993
; APPLICATION NUMBER: GB 9021290.3
; FILING DATE: 01-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9022570.7
; FILING DATE: 17-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9106546.6
; FILING DATE: 27-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 1498-30
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 176 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Salmonella enteritidis/Salmonella dublin
; US-08-449-922-1

Query Match 100.0%; Score 128; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 5.9e-115;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAQNTTSANWSDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGASVSGVATVPFV 60
DB 49 AAQNTTSANWSDPGFTGPAVAAGQKVGTLSTATGPHNSVSIAGKGASVSGVATVPFV 108

QY 61 DQGGQPVFRGRIQGANINDQANTGIDGLAGNRVASSQETLNVPTTFGKSTLPAGTFTAT 120
DB 109 DQGGQPVFRGRIQGANINDQANTGIDGLAGNRVASSQETLNVPTTFGKSTLPAGTFTAT 168

QY 121 FYVQQYQN 128
DB 169 FYVQQYQN 176

RESULT 4
US-09-230-078A-4
; Sequence 4, Application US/09230078A
; Patent No. 6495334
; GENERAL INFORMATION:
; APPLICANT: Rajashekara, Gireesh
; APPLICANT: Kakambi, Nagarata V.
; APPLICANT: Kapur, Vivek
; TITLE OF INVENTION: RECOMBINANT SEF14 FIMBRIAL PROTEIN FROM SALMONELLA
; FILE REFERENCE: 600.33USWO
; CURRENT APPLICATION NUMBER: US/09/230,078A
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: PCT/US97/12639
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: 60/022,191
; PRIOR FILING DATE: 1996-07-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Salmonella enteritidis
; US-09-230-078A-4

Query Match 100.0%; Score 128; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 6e-115;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAQNTTSANWSQDPGFTGPAVAGKVGKVTLSITATGPHNSVSIAGKGASVSGVATVPFV 60
Db 53 AAQNTTSANWSQDPGFTGPAVAGKVGKVTLSITATGPHNSVSIAGKGASVSGVATVPFV 112
QY 61 DQGGQPVFRGRIOGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 120
Db 113 DQGGQPVFRGRIOGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAGTFTAT 172

QY 121 FVVOQYQN 128

Db 173 FVVOQYQN 180

RESULT 5

US-09-543-407-46
; Sequence 46, Application US/09543407
; Patent No. 6864365
; GENERAL INFORMATION:
; APPLICANT: White, Aaron P.
; APPLICANT: Doran, James L.
; APPLICANT: Collinson, S. Karen
; APPLICANT: Kay, William W.
; TITLE OF INVENTION: BACTERIAL FIMBRIAL SYSTEM FOR
; TITLE OF INVENTION: PRESENTATION OF HETEROLOGOUS PEPTIDE SEQUENCES
; FILE REFERENCE: 920043.406
; CURRENT APPLICATION NUMBER: US/09/543,407
; CURRENT FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Salmonella enteritidis
US-09-543-407-46

Query Match 89.8%; Score 115; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.4e-102; Mismatches 0; Indels 0; Gaps 0;
Matches 115; Conservative 0;

QY 1 AAQNTTSANWSQDPGFTGPAVAGKVGKVTLSITATGPHNSVSIAGKGASVSGVATVPFV 60

Db 17 AAQNTTSANWSQDPGFTGPAVAGKVGKVTLSITATGPHNSVSIAGKGASVSGVATVPFV 76

QY 61 DQGGQPVFRGRIOGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAG 115

Db 77 DQGGQPVFRGRIOGANINDQANTGIDGLAGWRVASSQETLNVPVTTFGKSTLPAG 131

RESULT 6

US-09-688-017-194
; Sequence 194, Application US/09688017
; Patent No. 6942981
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Rabinowitz, Joshua D.
; APPLICANT: Schweizer, Johannes
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: Molecular Interactions in Hematopoietic
; TITLE OF INVENTION: Cells
; FILE REFERENCE: 020054-001110US
; CURRENT APPLICATION NUMBER: US/09/688,017
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/134,114
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,117
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,118
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453

; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 383
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 194
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: AA19L CD148 PL peptide
US-09-688-017-194

Query Match 5.5%; Score 7; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 PVTTFGK 109

Db 5 PVTTFGK 11

RESULT 7

US-09-688-017-217
; Sequence 217, Application US/09688017
; Patent No. 6942981
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Rabinowitz, Joshua D.
; APPLICANT: Schweizer, Johannes
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: Molecular Interactions in Hematopoietic
; TITLE OF INVENTION: Cells
; FILE REFERENCE: 020054-001110US
; CURRENT APPLICATION NUMBER: US/09/688,017
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/134,114
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,117
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,118
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 383
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 217

```
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: AA55L CD148 PL peptide
US-09-688-017-217

Query Match          5.5%; Score 7; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 PVTTFGK 109
Db 8 PVTTFGK 14

RESULT 8
US-09-443-184-53
; Sequence 53, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Mary Jane
; APPLICANT: Zweiger, Gary
; APPLICANT: Kaser, Matthew R.
; APPLICANT: Panzer, Scott
; APPLICANT: Seilhammer, Jeffrey J.
; APPLICANT: Yue, Henry
; APPLICANT: Baughn, Mariah
; APPLICANT: Azimzai, Valda
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
; FILE REFERENCE: PC-0007 US
; CURRENT APPLICATION NUMBER: US/09/443,184A
; CURRENT FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 53
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6372431 3123954CD1
US-09-443-184-53

Query Match          5.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 FGKSTLP 113
Db 60 FGKSTLP 66

RESULT 9
US-09-125-642C-15
; Sequence 15, Application US/09125642C
; Patent No. 6365393
; GENERAL INFORMATION:
; APPLICANT: BAYER AG
; TITLE OF INVENTION: Parapoxviruses Which Contain Foreign DNA, and
; Their Production and Their Use in Vaccines
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bayer Corporation
; CITY: Pittsburgh
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 15205-9741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/125,642C
; FILING DATE: 20-Aug-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/00729,
; FILING DATE: 17-Feb-97
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parapox ovis
; STRAIN: D1701 VEGF- Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-125-642C-15

Query Match          5.5%; Score 7; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GASVSGG 53
Db 91 GASVSGG 97

RESULT 10
US-09-431-888-11
; Sequence 11, Application US/09431888A
; Patent No. 6541008
; GENERAL INFORMATION:
; APPLICANT: Wise, Lyn M
; APPLICANT: Mercer, Andrew A
; APPLICANT: Savory, Loreen J
; APPLICANT: Fleming, Stephen B
; APPLICANT: Stacker, Stephen
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-LIKE PROTEIN FROM ORF
; TITLE OF INVENTION: VIRUS NZ2 BINDS AND ACTIVATES MAMMALIAN VEGF
; TITLE OF INVENTION: RECEPTOR-2, AND USES THEREOF
; FILE REFERENCE: Sequence Listing for 09/431,833
; Patent No. 6541008
; CURRENT APPLICATION NUMBER: US/09/431,888A
; CURRENT FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/106,689
; EARLIER FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: 60/106,800
; EARLIER FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Orf virus
US-09-431-888-11

Query Match          5.5%; Score 7; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GASVSGG 53
Db 91 GASVSGG 97

RESULT 11
US-08-470-179-28
; Sequence 28, Application US/08470179
```

; Patent No. 5645994
; GENERAL INFORMATION:
; APPLICANT: Huang Ph.D, Wai Mun
; TITLE OF INVENTION: Method and Compositions for
; TITLE OF INVENTION: Identification of Species in a Sample
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask, Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,179
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Ph.D, Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801-532-1922
; TELEFAX: 801-531-9168
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 143 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Legionella pneumoniae
US-08-470-179-28

Query Match 5.5%; Score 7; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 PFVDGQG 64
|||||
Db 64 PFVDGQG 70

RESULT 12
US-09-372-422A-42
; Sequence 42, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Zea mays
US-09-372-422A-42

Query Match 5.5%; Score 7; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GPAVAA 24
|||||
Db 189 GPAVAA 195

RESULT 13
US-09-902-540-13474
; Sequence 13474, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13474
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13474

Query Match 5.5%; Score 7; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TGPAAVAA 23
|||||
Db 195 TGPAAVAA 201

RESULT 14
US-09-372-422A-48
; Sequence 48, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Zea mays
US-09-372-422A-48

Query Match 5.5%; Score 7; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GPAVAA 24
|||||
Db 201 GPAVAA 207

RESULT 15
US-09-372-422A-22
; Sequence 22, Application US/09372422A

; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Zea mays
US-09-372-422A-22

Query Match 5.5%; Score 7; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GPAVAAG 24
| | | | |
Db 202 GPAVAAG 208

Search completed: April 26, 2006, 17:10:48
Job time : 48 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:10:03 ; Search time 117 Seconds
(without alignments)
457.112 Million cell updates/sec

Title: US-10-018-892-3

Perfect score: 128

Sequence: 1 AAQNTSANSQDPGFTGPA.....KSTLPAGTTFATFYVQYQN 128

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1867569 seqs, 41782326 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1864775

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query				DB	ID	Description
	Score	Match	Length				
1	9	7.0	606	4	US-10-156-761-9993	Sequence 9993, Ap	
2	8	6.2	194	4	US-10-437-963-150729	Sequence 150729,	
3	8	6.2	225	4	US-10-210-281-72	Sequence 72, Appl	
4	8	6.2	345	4	US-10-094-749-3021	Sequence 3021, Ap	
5	8	6.2	490	4	US-10-369-493-21753	Sequence 21753, A	
6	8	6.2	494	5	US-10-926-543-84	Sequence 84, Appl	
7	8	6.2	574	5	US-10-450-763-59630	Sequence 59630, A	
8	8	6.2	705	4	US-10-473-574-10	Sequence 10, Appl	
9	7	5.5	17	5	US-10-938-249-427	Sequence 427, App	
10	7	5.5	17	5	US-10-938-249-457	Sequence 457, App	
11	7	5.5	17	6	US-11-131-054-194	Sequence 194, App	
12	7	5.5	17	6	US-11-131-054-217	Sequence 217, App	
13	7	5.5	17	6	US-11-131-042-194	Sequence 194, App	
14	7	5.5	17	6	US-11-131-042-217	Sequence 217, App	
15	7	5.5	58	4	US-10-409-701-17	Sequence 17, Appl	
16	7	5.5	63	4	US-10-296-115-1437	Sequence 1437, Ap	
17	7	5.5	75	4	US-10-424-599-276075	Sequence 276075,	
18	7	5.5	78	4	US-10-425-115-238006	Sequence 238006,	
19	7	5.5	83	3	US-09-764-891-4562	Sequence 4562, Ap	
20	7	5.5	84	4	US-10-724-972A-6584	Sequence 6584, Ap	
21	7	5.5	84	4	US-10-425-115-320280	Sequence 320280,	
22	7	5.5	85	4	US-10-437-963-161218	Sequence 161218,	
23	7	5.5	104	4	US-10-317-832-54	Sequence 54, Appl	
24	7	5.5	104	5	US-10-733-878-54	Sequence 54, Appl	
25	7	5.5	104	5	US-10-601-072-54	Sequence 54, Appl	
26	7	5.5	114	3	US-09-890-688-66	Sequence 66, Appl	
27	7	5.5	114	4	US-10-333-900-20	Sequence 20, Appl	

28	7	5.5	121	4	US-10-767-701-54316	Sequence 54316, A
29	7	5.5	128	4	US-10-425-115-314419	Sequence 314419,
30	7	5.5	130	3	US-09-805-290A-24	Sequence 24, Appl
31	7	5.5	131	4	US-10-767-701-33366	Sequence 33366, A
32	7	5.5	132	3	US-09-795-006A-151	Sequence 151, App
33	7	5.5	132	4	US-10-352-153-11	Sequence 11, Appl
34	7	5.5	132	5	US-10-772-927A-30	Sequence 30, Appl
35	7	5.5	134	4	US-10-424-599-259533	Sequence 259533,
36	7	5.5	134	4	US-10-767-701-44321	Sequence 44321, A
37	7	5.5	139	3	US-09-925-298-673	Sequence 673, App
38	7	5.5	139	4	US-10-102-806-673	Sequence 673, App
39	7	5.5	143	4	US-10-437-963-147994	Sequence 147994,
40	7	5.5	171	4	US-10-424-599-164998	Sequence 164998,
41	7	5.5	174	4	US-10-282-122A-62343	Sequence 62343, A
42	7	5.5	179	4	US-10-424-599-216265	Sequence 216265,
43	7	5.5	196	4	US-10-437-963-115691	Sequence 115691,
44	7	5.5	196	5	US-10-450-763-33534	Sequence 33534, A
45	7	5.5	201	4	US-10-767-701-39996	Sequence 39996, A

ALIGNMENTS

RESULT 1
US-10-156-761-9993
; Sequence 9993, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, HIROSHI
; APPLICANT: HORIKAWA, JUN
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9993
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9993

Query Match 7.0%; Score 9; DB 4; Length 606;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 SVSGGVATV 57
DB 73 SVSGGVATV 81

RESULT 2
US-10-437-963-150729
; Sequence 150729, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 150729
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5093C.1.Dep
US-10-437-963-150729

Query Match          6.2%; Score 8; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      87 GLAGWRVA 94
Db      138 GLAGWRVA 145

RESULT 3
US-10-210-281-72
; Sequence 72, Application US/10210281
; Publication No. US20040030096A1
; GENERAL INFORMATION:
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Zhong, Mei
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Sciore, Paul
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Casman, Stacie
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-416D
; CURRENT APPLICATION NUMBER: US/10/210,281
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/361,775
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/361,832
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,203
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,702
; PRIOR FILING DATE: 2001-08-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191

```

```

; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 72
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-281-72

Query Match          6.2%; Score 8; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      23 AGQKVGTL 30
Db      16 AGQKVGTL 23

RESULT 4
US-10-094-749-3021
; Sequence 3021, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3021
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3021

Query Match          6.2%; Score 8; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      23 AGQKVGTL 30
Db      142 AGQKVGTL 149

RESULT 5
US-10-369-493-21753
; Sequence 21753, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

```



```
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21753
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Rhodobacter capsulatus
US-10-369-493-21753

Query Match          6.2%; Score 8; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AVAAGQKV 27
      |||||
Db 268 AVAAGQKV 275

RESULT 6
US-10-926-543-84
; Sequence 84, Application US/10926543
; Publication No. US2005004858A1
; GENERAL INFORMATION:
; APPLICANT: Jendoubi, Moncef
; TITLE OF INVENTION: LUNG CANCER SPECIFIC GENE PRODUCTS: THEIR CODING SEQUENCE, THEIR
; TITLE OF INVENTION: ANTIBODIES AND THEIR USE IN DIAGNOSTIC, THERAPEUTIC AND DISEASE
; TITLE OF INVENTION: MANAGEMENT OF LUNG CANCER
; FILE REFERENCE: 705403.4004
; CURRENT APPLICATION NUMBER: US/10/926,543
; CURRENT FILING DATE: 2004-08-25
; PRIOR APPLICATION NUMBER: US 60/497,790
; PRIOR FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 84
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-926-543-84

Query Match          6.2%; Score 8; DB 5; Length 494;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AGQKVGT 30
      |||||
Db 185 AGQKVGT 192

RESULT 7
US-10-450-763-59630
; Sequence 59630, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyveeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
```

```
; SEQ ID NO 59630
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (326)..(351)
; OTHER INFORMATION: CAP-Gly domain proteins domain identified by eMATRIX,
; OTHER INFORMATION: accession number BL00845, p-value=9.820e-19, raw score of 16.4
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (319)..(367)
; OTHER INFORMATION: CAP-Gly domain identified by Pfam, accession name CAP_GLY, E-
; OTHER INFORMATION: value=1.1e-15, Pfam score of 65.6
US-10-450-763-59630

Query Match          6.2%; Score 8; DB 5; Length 574;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AGQKVGT 30
      |||||
Db 308 AGQKVGT 315

RESULT 8
US-10-473-574-10
; Sequence 10, Application US/10473574
; Publication No. US20040116670A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; HAPALIA, April J.A.;
; APPLICANT: TANG, Y. Tom; YUE, Henry;
; APPLICANT: KHAN, Farrah A.; ISON, Craig H.;
; APPLICANT: BAUGHN, Mariah R.; WARREN, Bridget A.;
; APPLICANT: DUGGAN, Brendan M.; THANGAVELU, Kavitha;
; APPLICANT: HONCHELL, Cynthia D.; AZIMZAI, Yalda;
; APPLICANT: ELLIOTT, Vicki S.; BURFORD, Neil;
; APPLICANT: DING, Li; YUE, Huibin;
; APPLICANT: BECHA, Shanya; EMERLING, Brooke M.;
; APPLICANT: RICHARDSON, Thomas W.; LEE, Soo Yeun;
; APPLICANT: BANDMAN, Olga; LAL, Preeti G.;
; APPLICANT: LEE, Sally; GIETZEN, Kimberly J.;
; APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
; APPLICANT: LEE, Ernestine A.; SWARNAKAR, Anita;
; APPLICANT: KING, Huijun Z.; JONES, Karen Anne
; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0918 USN
; CURRENT APPLICATION NUMBER: US/10/473,574
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: PCT/US02/09288
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US 60/294,451
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/291,870
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/290,518
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/288,609
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/283,769
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/281,323
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/280,508
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

OTHER INFORMATION: Incyte ID No: 1683662CD1
US-10-473-574-10

Query Match 6.2%; Score 8; DB 4; Length 705;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AGQKVGTL 30
DB 292 AGQKVGTL 299

RESULT 9
US-10-938-249-427
; Sequence 427, Application US/10938249
; Publication No. US20050037969A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Rabinowitz, Joshua D.
; APPLICANT: Schweizer, Johannes
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: Molecular Interactions in Hematopoietic
; FILE REFERENCE: 020054-001130US
; CURRENT APPLICATION NUMBER: US/10/938,249
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US/09/724,553
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/134,114
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,117
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,118
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,117
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 543
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 427
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PL peptide
US-10-938-249-427

Query Match 5.5%; Score 7; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 PVTTFGK 109
DB 5 PVTTFGK 11

RESULT 10
US-10-938-249-457
; Sequence 457, Application US/10938249
; Publication No. US20050037969A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Rabinowitz, Joshua D.

APPLICANT: Schweizer, Johannes
APPLICANT: Arbor Vita Corporation
TITLE OF INVENTION: Molecular Interactions in Hematopoietic
FILE REFERENCE: 020054-001130US
CURRENT APPLICATION NUMBER: US/10/938,249
CURRENT FILING DATE: 2004-09-10
PRIOR APPLICATION NUMBER: US/09/724,553
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/134,114
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: US 60/134,117
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: US 60/134,118
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: US 60/160,860
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: US 60/162,498
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: US 60/170,453
PRIOR FILING DATE: 1999-12-13
PRIOR APPLICATION NUMBER: US 60/176,195
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: US 60/182,296
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: US 60/196,267
PRIOR FILING DATE: 2000-04-11
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 543
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 457
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PL peptide
US-10-938-249-457

Query Match 5.5%; Score 7; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 PVTTFGK 109
DB 8 PVTTFGK 14

RESULT 11
US-11-131-054-194
; Sequence 194, Application US/11131054
; Publication No. US20050214869A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Rabinowitz, Joshua D.
; APPLICANT: Schweizer, Johannes
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: Molecular Interactions in Hematopoietic
; FILE REFERENCE: 020054-001110US
; CURRENT APPLICATION NUMBER: US/11/131,054
; CURRENT FILING DATE: 2005-05-16
; PRIOR APPLICATION NUMBER: US/09/688,017
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/134,114
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,117
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,118
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29

;; PRIOR APPLICATION NUMBER: US 60/170,453
;; PRIOR FILING DATE: 1999-12-13
;; PRIOR APPLICATION NUMBER: US 60/176,195
;; PRIOR FILING DATE: 2000-01-14
;; PRIOR APPLICATION NUMBER: US 60/182,296
;; PRIOR FILING DATE: 2000-02-14
;; PRIOR APPLICATION NUMBER: US 60/196,267
;; PRIOR FILING DATE: 2000-04-11
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 383
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 194
;; LENGTH: 17
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: AA19L CD148 PL peptide
US-11-131-054-194

Query Match 5.5%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 PVTTFGK 109
Db 5 PVTTFGK 11

RESULT 12
US-11-131-054-217
;; Sequence 217, Application US/11131054
;; Publication No. US20050214869A1
;; GENERAL INFORMATION:
;; APPLICANT: Lu, Peter S.
;; APPLICANT: Rabinowitz, Joshua D.
;; APPLICANT: Schweizer, Johannes
;; APPLICANT: Arbor Vita Corporation
;; TITLE OF INVENTION: Molecular Interactions in Hematopoietic
;; TITLE OF INVENTION: Cells
;; FILE REFERENCE: 020054-001110US
;; CURRENT APPLICATION NUMBER: US/11/131,054
;; CURRENT FILING DATE: 2005-05-16
;; PRIOR APPLICATION NUMBER: US/09/688,017
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/134,114
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: US 60/134,118
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: US 60/160,860
;; PRIOR FILING DATE: 1999-10-21
;; PRIOR APPLICATION NUMBER: US 60/162,498
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: US 60/170,453
;; PRIOR FILING DATE: 1999-12-13
;; PRIOR APPLICATION NUMBER: US 60/176,195
;; PRIOR FILING DATE: 2000-01-14
;; PRIOR APPLICATION NUMBER: US 60/182,296
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: US 60/170,453
;; PRIOR FILING DATE: 1999-12-13
;; PRIOR APPLICATION NUMBER: US 60/176,195
;; PRIOR FILING DATE: 2000-01-14
;; PRIOR APPLICATION NUMBER: US 60/182,296
;; PRIOR FILING DATE: 2000-02-14
;; PRIOR APPLICATION NUMBER: US 60/196,267
;; PRIOR FILING DATE: 2000-04-11
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 383
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 217
;; LENGTH: 17
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: AA55L CD148 PL peptide
US-11-131-054-217

Query Match 5.5%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 PVTTFGK 109
Db 8 PVTTFGK 14

RESULT 13
US-11-131-042-194
;; Sequence 194, Application US/11131042
;; Publication No. US20050221388A1
;; GENERAL INFORMATION:
;; APPLICANT: Lu, Peter S.
;; APPLICANT: Rabinowitz, Joshua D.
;; APPLICANT: Schweizer, Johannes
;; APPLICANT: Arbor Vita Corporation
;; TITLE OF INVENTION: Molecular Interactions in Hematopoietic
;; TITLE OF INVENTION: Cells
;; FILE REFERENCE: 020054-001110US
;; CURRENT APPLICATION NUMBER: US/11/131,042
;; CURRENT FILING DATE: 2005-05-16
;; PRIOR APPLICATION NUMBER: US/09/688,017
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: US 60/134,114
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: US 60/134,117
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: US 60/134,118
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: US 60/160,860
;; PRIOR FILING DATE: 1999-10-21
;; PRIOR APPLICATION NUMBER: US 60/162,498
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: US 60/170,453
;; PRIOR FILING DATE: 1999-12-13
;; PRIOR APPLICATION NUMBER: US 60/176,195
;; PRIOR FILING DATE: 2000-01-14
;; PRIOR APPLICATION NUMBER: US 60/182,296
;; PRIOR FILING DATE: 2000-02-14
;; PRIOR APPLICATION NUMBER: US 60/196,267
;; PRIOR FILING DATE: 2000-04-11
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 383
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 194
;; LENGTH: 17
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: AA19L CD148 PL peptide
US-11-131-042-194

Query Match 5.5%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 PVTTFGK 109
Db 5 PVTTFGK 11

RESULT 14
US-11-131-042-217
;; Sequence 217, Application US/11131042
;; Publication No. US20050221388A1
;; GENERAL INFORMATION:
;; APPLICANT: Lu, Peter S.
;; APPLICANT: Rabinowitz, Joshua D.
;; APPLICANT: Schweizer, Johannes
;; APPLICANT: Arbor Vita Corporation
;; TITLE OF INVENTION: Molecular Interactions in Hematopoietic

Search completed: April 26, 2006, 17:12:57
Job time : 118 secs

```

; TITLE OF INVENTION: Cells
; FILE REFERENCE: 020054-001110US
; CURRENT APPLICATION NUMBER: US/11/131,042
; CURRENT FILING DATE: 2005-05-16
; PRIOR APPLICATION NUMBER: US/09/688,017
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/134,114
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,117
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/134,118
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 217
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: AA55L CD148 PL peptide
US-11-131-042-217

Query Match      5.5%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 PVTTFGK 109
Db 8 PVTTFGK 14

RESULT 15
US-10-409-701-17
; Sequence 17, Application US/10409701
; Publication No. US20030221224A1
; GENERAL INFORMATION:
; APPLICANT: Zinselmeier, Chris
; TITLE OF INVENTION: Enhanced Silk Exsersion Under Stress
; FILE REFERENCE: 1421
; CURRENT APPLICATION NUMBER: US/10/409,701
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/370,796
; PRIOR FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Zea mays
US-10-409-701-17

Query Match      5.5%; Score 7; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GPAVAAG 24
Db 8 GPAVAAG 14

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:10:00 ; Search time 13 Seconds
(without alignments)
447.782 Million cell updates/sec

Title: US-10-018-892-3
Perfect score: 128
Sequence: 1 AAQNTTSANMSQDPFGTGA.....KSTLPAGTTFATFVQYQVN 128

Scoring table: Oligo

Gapop 60.0 , Gapext 60.0

Searched: 232119 seqs, 45477862 residues

Word size: 1

Total number of hits satisfying chosen parameters: 231540

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA New:
1: /SID55/ptodata/2/pubpaa/US08 NEW PUB.pep.*
2: /SID55/ptodata/2/pubpaa/US06 NEW PUB.pep.*
3: /SID55/ptodata/2/pubpaa/US07 NEW PUB.pep.*
4: /SID55/ptodata/2/pubpaa/PTC NEW PUB.pep.*
5: /SID55/ptodata/2/pubpaa/US09 NEW PUB.pep.*
6: /SID55/ptodata/2/pubpaa/US10 NEW PUB.pep.*
7: /SID55/ptodata/2/pubpaa/US11 NEW PUB.pep.*
8: /SID55/ptodata/2/pubpaa/US60 NEW PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	5.5	132	7 US-11-073-605-15	Sequence 15, Appl
2	7	5.5	132	7 US-11-064-774A-151	Sequence 151, Appl
3	7	5.5	132	7 US-11-075-400-22	Sequence 22, Appl
4	7	5.5	132	7 US-11-075-047A-93	Sequence 93, Appl
5	7	5.5	137	6 US-10-995-561-853	Sequence 853, Appl
6	7	5.5	247	7 US-11-087-099-5549	Sequence 5549, Appl
7	7	5.5	292	7 US-11-087-099-9937	Sequence 9937, Appl
8	7	5.5	329	7 US-11-096-568A-347	Sequence 347, Appl
9	7	5.5	331	7 US-11-096-568A-346	Sequence 346, Appl
10	7	5.5	420	7 US-11-188-298-15565	Sequence 15565, A
11	7	5.5	454	7 US-11-045-004-1402	Sequence 1402, Appl
12	7	5.5	479	7 US-11-188-298-12225	Sequence 12225, A
13	7	5.5	497	7 US-11-188-298-7036	Sequence 7036, Appl
14	7	5.5	497	7 US-11-188-298-13926	Sequence 13926, A
15	7	5.5	497	7 US-11-188-298-15143	Sequence 15143, A
16	7	5.5	497	7 US-11-188-298-21056	Sequence 21056, A
17	7	5.5	501	7 US-11-094-917-36	Sequence 36, Appl
18	7	5.5	555	7 US-11-096-568A-21949	Sequence 21949, A
19	7	5.5	802	7 US-11-045-004-1971	Sequence 1971, Appl
20	7	5.5	827	7 US-11-079-463-5812	Sequence 5812, Appl
21	7	5.5	1075	7 US-11-079-463-6009	Sequence 6009, Appl
22	7	5.5	1178	6 US-10-995-561-851	Sequence 851, Appl
23	7	5.5	1337	7 US-11-112-304A-33	Sequence 33, Appl
24	7	5.5	2117	7 US-11-087-099-9594	Sequence 9594, Appl
25	7	5.5	2117	7 US-11-188-298-8910	Sequence 8910, Appl

ALIGNMENTS

RESULT 1

US-11-073-605-15
; Sequence 15, Application US/11073605
; Publication No. US20050260161A1
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, ULF
; APPLICANT: ALITALO, KARI
; APPLICANT: UUTELA, MARKO
; TITLE OF INVENTION: METHOD FOR MODULATING REGULATING AND/OR STABILIZING
; FILE REFERENCE: 029065.53347US
; CURRENT APPLICATION NUMBER: US/11/073,605
; CURRENT FILING DATE: 2005-03-08
; PRIOR APPLICATION NUMBER: 60/550,327
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 15
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Orf virus
US-11-073-605-15

Query Match 5.5%; Score 7; DB 7; Length 132;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 GASVSGG 53
Db 91 GASVSGG 97

RESULT 2

US-11-064-774A-151
; Sequence 151, Application US/11064774A
; Publication No. US20050267024A1
; GENERAL INFORMATION:
; APPLICANT: Alicata et al.
; TITLE OF INVENTION: FACTOR DNAs AND PROTEINS
; FILE REFERENCE: 28967/3597B2
; CURRENT APPLICATION NUMBER: US/11/064,774A
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: 09/795,006
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 60/205,331
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/185,205

Sequence 2434, Ap
Sequence 2344, Ap
Sequence 819, App
Sequence 5017, Ap
Sequence 6714, Ap
Sequence 7862, Ap
Sequence 2066, Ap
Sequence 24144, A
Sequence 1508, Ap
Sequence 9556, Ap
Sequence 1495, Ap
Sequence 1453, A
Sequence 16767, A
Sequence 1294, Ap
Sequence 24143, A
Sequence 38, Appl
Sequence 6, Appl
Sequence 37, Appl
Sequence 508, App
Sequence 16766, A

;; PRIOR FILING DATE: 2000-02-25
;; NUMBER OF SEQ ID NOS: 1212
;; SOFTWARE: PatentIn ver. 3.2
;; SEQ ID NO 151
;; LENGTH: 132
;; TYPE: PRT
;; ORGANISM: Orf virus
US-11-064-774A-151

Query Match 5.5%; Score 7; DB 7; Length 132;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GASVSGG 53
Db 91 GASVSGG 97

RESULT 3

US-11-075-400-22
; Sequence 22, Application US/11075400
; Publication No. US2005028223A1
; GENERAL INFORMATION:
; APPLICANT: ERIKSON, et al.
; TITLE OF INVENTION: MULTIVALENT ANTIBODY MATERIALS AND METHODS FOR VEGF/PDGF FAMILY OR
; TITLE OF INVENTION: GROWTH FACTORS
; FILE REFERENCE: 28967/39820B
; CURRENT APPLICATION NUMBER: US/11/075,400
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/550,511
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/586,662
; PRIOR FILING DATE: 2004-07-09
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 132
; TYPE: PRT
; ORGANISM: ORF Virus
US-11-075-400-22

Query Match 5.5%; Score 7; DB 7; Length 132;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GASVSGG 53
Db 91 GASVSGG 97

RESULT 4

US-11-075-047A-93
; Sequence 93, Application US/11075047A
; Publication No. US20060030000A1
; GENERAL INFORMATION:
; APPLICANT: ALITALO, et al.
; TITLE OF INVENTION: GROWTH FACTOR BINDING CONSTRUCTS MATERIALS AND METHODS
; FILE REFERENCE: 28967/39700A
; CURRENT APPLICATION NUMBER: US/11/075,047A
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/550,907
; PRIOR FILING DATE: 2004-03-07
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 93
; LENGTH: 132
; TYPE: PRT
; ORGANISM: ORF virus
US-11-075-047A-93

Query Match 5.5%; Score 7; DB 7; Length 132;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GASVSGG 53
Db 91 GASVSGG 97

RESULT 5

US-10-995-561-853
; Sequence 853, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 853
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-853

Query Match 5.5%; Score 7; DB 6; Length 137;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 PVTTFGK 109
Db 49 PVTTFGK 55

RESULT 6

US-11-087-099-5549
; Sequence 5549, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5549
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-11-087-099-5549

Query Match 5.5%; Score 7; DB 7; Length 247;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TGPAAVA 23
Db 197 TGPAAVA 203

RESULT 7

US-11-087-099-9937
; Sequence 9937, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9937

```
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Caulobacter crescentus CB15
US-11-087-099-9937

Query Match      5.5%; Score 7; DB 7; Length 292;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 IAGKGAS 49
Db 64 IAGKGAS 70

RESULT 8
US-11-096-568A-347
; Sequence 347, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 347
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(329)
; OTHER INFORMATION: Ceres Seq. ID no. 15180143
US-11-096-568A-347

Query Match      5.5%; Score 7; DB 7; Length 329;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 PHNSVSI 43
Db 117 PHNSVSI 123

RESULT 9
US-11-096-568A-346
; Sequence 346, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 346
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(331)
; OTHER INFORMATION: Ceres Seq. ID no. 15180142
US-11-096-568A-346

Query Match      5.5%; Score 7; DB 7; Length 331;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 PHNSVSI 43
Db 117 PHNSVSI 123

; LENGTH: 292
; TYPE: PRT
; ORGANISM: Caulobacter crescentus CB15
US-11-087-099-9937

Query Match      5.5%; Score 7; DB 7; Length 292;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 IAGKGAS 49
Db 64 IAGKGAS 70

RESULT 10
US-11-188-298-15565
; Sequence 15565, Application US/11188298
; Publication No. US2006007552A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15565
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Haemophilus somnus 129PT
US-11-188-298-15565

Query Match      5.5%; Score 7; DB 7; Length 420;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 TLPAGTF 117
Db 400 TLPAGTF 406

RESULT 11
US-11-045-004-1402
; Sequence 1402, Application US/11045004
; Publication No. US20060078901A1
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANCK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBELI, WERNER
; APPLICANT: KREFT, JURGEN
; APPLICANT: KUHN, MICHAEL
; APPLICANT: NG, EVA
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
; APPLICANT: GARRIDO-GARCIA, PATRICIA
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
; APPLICANT: AMEND, ALEXANDRA
; APPLICANT: CHAKRABORTY, TRINAD
; APPLICANT: DOMANN, EUGEN
; APPLICANT: HAIN, THORSTEN
; APPLICANT: BERCHE, PATRICK
; APPLICANT: CHARBIT, ALAIN
; APPLICANT: DURANT, LIONEL
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
; APPLICANT: BAQUERO, FERNANDO
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
; APPLICANT: GOMEZ-LOPEZ, NURIA
; APPLICANT: MADUENIO, ENCARNA
; APPLICANT: PABLOS, BETRIZ DE
; APPLICANT: WEHLAND, JURGEN
; APPLICANT: KARST, UWE
; APPLICANT: ENTIAN, KARL-DIETER
```

```

; APPLICANT: HAUF, JORG
; APPLICANT: ROSE, MATTHIAS
; APPLICANT: VOSS, HAMUT
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1402
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
; US-11-045-004-1402

Query Match      5.5%; Score 7; DB 7; Length 454;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      110 STLPGT 116
Db      362 STLPGT 368

RESULT 12
US-11-188-298-12225
; Sequence 12225, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 12225
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor A3(2)
; US-11-188-298-12225

Query Match      5.5%; Score 7; DB 7; Length 479;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GPAVAA 24
Db      163 GPAVAA 169

RESULT 13
US-11-188-298-7036
; Sequence 7036, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569

```

```

; SEQ ID NO 7036
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Emericella nidulans
; US-11-188-298-7036

Query Match      5.5%; Score 7; DB 7; Length 497;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GPAVAA 24
Db      176 GPAVAA 182

RESULT 14
US-11-188-298-13926
; Sequence 13926, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 13926
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Emericella nidulans
; US-11-188-298-13926

Query Match      5.5%; Score 7; DB 7; Length 497;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GPAVAA 24
Db      176 GPAVAA 182

RESULT 15
US-11-188-298-15143
; Sequence 15143, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15143
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Emericella nidulans
; US-11-188-298-15143

Query Match      5.5%; Score 7; DB 7; Length 497;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GPAVAA 24
Db      176 GPAVAA 182

RESULT 16
US-11-188-298-15143
; Sequence 15143, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15143
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Emericella nidulans
; US-11-188-298-15143

Query Match      5.5%; Score 7; DB 7; Length 497;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GPAVAA 24
Db      176 GPAVAA 182

Search completed: April 26, 2006, 17:10:21
Job time : 14 secs

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2006, 17:03:33 ; Search time 15 Seconds
(without alignments)
821.050 Million cell updates/sec

Title: US-10-018-892-3

Perfect score: 662

Sequence: 1 AAQNTTSANWSQDPGFTGPA.....KSTLPAAGTTATFYVQYQN 128

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	662	100.0	165	2	A40618	fimbrin, SEF14 - S
2	91.5	13.8	319	2	T43040	hypothetical prote
3	91.5	13.8	2468	2	A83412	hypothetical prote
4	87.5	13.2	380	1	GNV5MB	genome polyprotein
5	85	12.8	786	2	T16509	hypothetical prote
6	84	12.7	591	2	AC3528	extracellular seri
7	83	12.5	549	2	B70508	probable htrA prot
8	83	12.5	1844	2	T51890	related to Nup98-N
9	82.5	12.5	1045	2	A39199	endoglucanase B (E
10	81.5	12.3	674	2	B82954	conserved hypothet
11	81.5	12.3	1373	1	A43291	collagen alpha 2(I
12	81.5	12.3	2204	2	A70524	probable PPS prote
13	79.5	12.0	1286	2	S28634	adhesin AIDA-I pre
14	79.5	12.0	1509	2	B89985	hypothetical prote
15	78.5	11.9	320	2	G97016	comE-like protein,
16	78.5	11.9	645	1	SUBSMP	serine proteinase
17	78.5	11.9	1417	2	A83080	hypothetical prote
18	78.5	11.9	1615	2	B49502	protein-tyrosine-p
19	78.5	11.9	1651	2	JC1340	outer membrane pro
20	78.5	11.9	1767	2	A49502	protein-tyrosine-p
21	78	11.8	445	2	T34352	hypothetical prote
22	78	11.8	751	2	SE4741	cuticle collagen -
23	77.5	11.7	180	2	AC0208	probable fimbrin
24	77.5	11.7	1109	2	A56143	surface-array prot
25	77.5	11.7	1665	2	T29008	hypothetical prote
26	77	11.6	469	2	A35789	glutamate-1-semial
27	77	11.6	1024	2	H87599	TonB-dependent rec
28	77	11.6	1569	2	A65044	hypothetical prote
29	76.5	11.6	255	2	T50916	hypothetical prote

30	76.5	11.6	713	1	UMMS	period clock prote
31	76.5	11.6	792	1	EAHU	elastin precursor,
32	76.5	11.6	1335	2	G90975	probable factor [i
33	76.5	11.6	2383	2	D64962	probable membrane
34	76.5	11.6	2660	2	E85822	probable invasin Z
35	76	11.5	327	2	C97057	spore protease [im
36	76	11.5	395	2	AG2606	conserved hypothet
37	76	11.5	395	2	F97388	probable lipase (A
38	76	11.5	1052	2	AF2959	conserved hypothet
39	76	11.5	1219	2	T14578	nucleoporin Nup153
40	76	11.5	1341	2	H98323	hypothetical prote
41	75.5	11.4	309	1	KIECRB	ribokinase (SC 2.7
42	75.5	11.4	309	1	F91215	ribokinase [import
43	75.5	11.4	309	2	G86061	ribokinase [import
44	75.5	11.4	673	1	VCPVB5	coat protein VP1 -
45	75.5	11.4	1145	2	B75625	hypothetical prote

ALIGNMENTS

RESULT 1

A40618
fimbrin, SEF14 - Salmonella enteritidis
N:Alternate names: fimbrin, Sefa
C:Species: Salmonella enteritidis
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A40618; A25034
R:Clouthier, S.C.; Muller, K.H.; Doran, J.L.; Collinson, S.K.; Kay, W.W.
J. Bacteriol. 175, 2523-2533, 1993
A:Title: Characterization of three fimbrin genes, sefABC, of Salmonella enteritidis.
A:Reference number: A40618; MUID:93239677; PMID:8097515
A:Contents: 27655-3b
A:Accession: A40618
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-165 <CLO>
A:Cross-references: UNIPROT:P12061; UNIPARC:UPI00001135BD; GB:L11008; NID:g310645; PIR:
A>Note: sequence extracted from NCBI backbone (NCBI:130387, NCBIP:130395)
R:Feutrier, J.; Kay, W.W.; Trust, T.J.
J. Bacteriol. 168, 221-227, 1986
A:Title: Purification and characterization of fimbriae from Salmonella enteritidis.
A:Reference number: A25034; MUID:87008384; PMID:2875990
A:Accession: A25034
A:Molecule type: protein
A:Residues: 22-78, 'X', 80-82, 'XQ', 128 <PFU>
A:Cross-references: UNIPARC:UPI0000179BD8

Query Match 100.0%; Score 662; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.6e-50; Indels 0; Gaps 0;
Matches 128; Conservative 0; Mismatches 0;

QY	1	AAQNTTSANWSQDPGFTGPAAGQKVGTLSTATGPHNSVSIAGKASVSGGVATVPFV	60
DB	38	AAQNTTSANWSQDPGFTGPAAGQKVGTLSTATGPHNSVSIAGKASVSGGVATVPFV	97
QY	61	DQCGQPFVRGRITQGANINDQANTGIDGLAGWRVASSQETLNVPVTFGKSTLPAGTTAT	120
DB	98	DQCGQPFVRGRITQGANINDQANTGIDGLAGWRVASSQETLNVPVTFGKSTLPAGTTAT	157
QY	121	FYVQYQYN	128
DB	158	FYVQYQYN	165

RESULT 2

T43040
hypothetical protein - fission yeast (Schizosaccharomyces pombe) (fragment)
C:Species: Schizosaccharomyces pombe
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43040
R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997

A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.

A;Reference number: Z17233; MUID:98162722; PMID:9501991
A;Accession: T43040
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-319 <YOS>
A;Cross-references: UNIPROT:Q8TFG9; UNIPARC:UPI00000698D1; EMBL:D89219; NID:gi749645; P1
A;Experimental source: strain PR745

Query Match 13.8%; Score 91.5; DB 2; Length 319;
Best Local Similarity 23.5%; Pred. No. 1.1;
Matches 36; Conservative 15; Mismatches 57; Indels 45; Gaps 5;

QY 4 NTTSANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKGSVSGGVATV-----57

DB 120 NPTAAPNE-----PVTVTGTEGSGVTTETPTTSTFSFTTIIIGTTIIPVPG 173

QY 58 -----PFVDGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSQETLNV 102

DB 174 NPSSVSAPPTTSFTPGGSGYPYSNTTQGN-----TTSI-----WNSSNSTIVSNV 223

QY 103 PVTTFGKSTLPAG-----TTFATFYVOOYQ 127

DB 224 TAITTGNVTITGDLTITDPTTFTSTYLSGFGQ 256

RESULT 3

A83412
hypothetical protein PA1874 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83412
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83412
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2468 <STO>
A;Cross-references: UNIPROT:Q912M3; UNIPARC:UPI00000C54E3; GB:AE004613; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1874

Query Match 13.8%; Score 91.5; DB 2; Length 2468;
Best Local Similarity 26.3%; Pred. No. 10;
Matches 36; Conservative 11; Mismatches 43; Indels 47; Gaps 7;

QY 7 SANWS-----QDP-GFTGPAVAAGQKVGTLISITATGPHNSVSIAGKA 48

DB 1284 SGNSFTPGTPLANGTVVNAVAQDPAGNTGQ-----GSTTVDVAVPTPVVNPNGN 1336

QY 49 SVSGGV---ATVPFVDGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSQETLNV---V 102

DB 1337 LLNGTABPGSTVTLTDGNGNPI-----GQTADGSGNWSFTPGSQLENGTVV 1383

QY 103 PVT---TFGKSTLPAGT 116

DB 1384 NVTASDAAGNTSLPAT 1400

RESULT 4

GNVSNB
Genome polyprotein - maize dwarf mosaic virus (strain B) (fragment)
N;Contains: carboxyl end of nuclear inclusion protein b; coat protein
C;Species: maize dwarf mosaic virus, MDMV
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C;Accession: PH0208
R;Frenkel, M.J.; Jilka, J.M.; McKern, N.M.; Strike, P.M.; Clark Jr., J.M.; Shukla, D.D.;
J. Gen. Virol. 72, 237-242, 1991

A;Title: Unexpected sequence diversity in the amino-terminal ends of the coat proteins
A;Reference number: PH0207; MUID:91132116; PMID:1993866
A;Accession: PH0208
A;Molecule type: genomic RNA

A;Residues: 1-380 <FRE>
A;Cross-references: UNIPROT:P32652; UNIPARC:UPI0000131E4F; GB:D00949; NID:g222059; P1D
C;Superfamily: tobacco etch virus genome polyprotein
C;Keywords: coat protein; inclusion protein
P;1-52/Product: nuclear inclusion protein b (fragment) #status predicted <NIP>
P;53-380/Product: coat protein #status predicted <COA>

Query Match 13.2%; Score 87.5; DB 1; Length 380;
Best Local Similarity 30.1%; Pred. No. 2.9;
Matches 31; Conservative 9; Mismatches 54; Indels 9; Gaps 2;

QY 11 SQDQFTGPAVAAGQKVGTLISITATGPHNSVSIAGKGSVSGGVATVPFVDGQGPVFRG 70

DB 64 SGSQGTTPATGSGAK-----PATSGAGSGGTGAGTGTGQARTSGTGTGSGATGG 117

QY 71 RIQGANINDQANTGIDGLAGRWVASSQETLNVFVTTFGKSTLP 113

DB 118 QSGSGSGTEQVNT---GSAGTNATGQDRDQDVGSTGKISVP 157

RESULT 5

Ti6509
hypothetical protein F59A6.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: Ti6509
R;Nhan, M.
submitted to the EMBL Data Library, December 1995
A;Description: The sequence of C. elegans cosmid F59A6.
A;Reference number: Z18526
A;Accession: Ti6509
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-786 <NHA>
A;Cross-references: UNIPROT:Q21027; UNIPARC:UPI000007A7BF; EMBL:U41994; NID:gl123047;
C;Genetics:
A;Gene: CESP:F59A6.3
A;Introns: 106/1; 620/1; 634/3; 675/3; 775/1

Query Match 12.8%; Score 85; DB 2; Length 786;
Best Local Similarity 27.4%; Pred. No. 11;
Matches 34; Conservative 15; Mismatches 63; Indels 12; Gaps 5;

QY 4 NTTS--ANWSQDPGFTGPAVAAGQKVGTLISITATGPHNSVSIAGKGSVSGGVATVP---58

DB 348 STTSRASSASDDPTTGTGPTTSGSTASTTS---GSLFSTSL-GSSQSPGSSVSTTPGPS 402

QY 59 FVDGQGPVFRGRIQGANINDQANTGIDGLAGRWVASSQETL--NVPVTFGKSTLPAGT 116

DB 403 TTSGISQSTSGTPTTTSBPTTSGTSDTSGPSTTSGPSTTTLGTTQSTTSGPSTTPGST 462

QY 117 FTAT 120

DB 463 ISTT 466

RESULT 6

AC3528
extracellular serine proteinase (EC 3.4.21.-) [imported] - Brucella melitensis (strain
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AC3528
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lel
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3528
A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-591 <KUR>
A:Cross-references: UNIPROT:Q8YDM6; UNIPARC:UPI0000058384; GB:AE008918; PIDN:AAL53390.1
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110149
A:Map position: 11
C:Keywords: hydrolase; serine proteinase

Query Match 12.7%; Score 84; DB 2; Length 591;
Best Local Similarity 31.4%; Pred. No. 9.4;
Matches 37; Conservative 17; Mismatches 38; Indels 26; Gaps 8;
QY 6 TSANWSQDPGFTGPAVAAGQKVTLSITAT----GPHNSVSIA-KGASVSGVA--TVP 58
DB 129 TGAN-----TPSGNIIVTG---GTLNVSSAALGAANEISLANGAGLSSGSLAGRSVT 180
QY 59 FVDGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTF--GKSTLPAG 115
DB 181 LTGGQ-----AAIGGAGVGDHFTGAGGLR-----ASSVTLSDSDNDYTGQTSLSG 228

RESULT 7
B70508
probable htrA protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: B70508
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70508
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-549 <COL>
A:Cross-references: UNIPROT:O06291; UNIPARC:UPI00000C09DE; GB:Z98260; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: htrA
C:Superfamily: Escherichia coli trypsin-like proteinase degS; GLGF domain homology; tryp

Query Match 12.5%; Score 83; DB 2; Length 549;
Best Local Similarity 25.6%; Pred. No. 11;
Matches 34; Conservative 18; Mismatches 43; Indels 38; Gaps 5;
QY 1 AAQNTTSANWSQDPGFTGPAVAAGQKVTLSITATGPHNSVSIAK-----GASVS- 51
DB 142 AAEKDGAGDFDPDPWRDPAAAA--LGTPALAAPAPGALGAGSGKLGVRDVLFGGKVS 199
QY 52 ---GGVATVPVVDGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQE---TLNVPVT 105
DB 200 LALGILVAIALVIG-----GIGGVIRKRTAEVVDATTSKVTL 238
QY 106 TFGKSTLPAGTFT 118
DB 239 TTGNAQEPAGRT 251

RESULT 8
T51890
related to Nup98-Nup96 precursor [imported] - Neurospora crassa
N:Alternate names: protein B23111.20
C:Species: Neurospora crassa
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T51890
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, August 2000
A:Reference number: 225858
A:Accession: T51890

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1844 <SCH>
A:Cross-references: UNIPARC:UPI000017B4PF; EMBL:AL391572; GSPDB:GN00116; NCSP:B23111
A:Experimental source: BAC clone B23111; strain OR74A
C:Genetics:
A:Gene: NCSP:B23111.20
A:Map position: 6
A:Introns: 34/1; 1281/3

Query Match 12.5%; Score 83; DB 2; Length 1844;
Best Local Similarity 30.8%; Pred. No. 41;
Matches 40; Conservative 9; Mismatches 55; Indels 26; Gaps 8;
QY 5 TTSANWSQDPGFTG-----PAVAAGQKVTLSITATGPHNSVSIAKGASVSGVATV 57
DB 356 TTTTGFGQNTGTTGGGLFGQQAAPAGFGFTTCAT-TGTGTSTG-TGFGSGTTGGFGTT 413
QY 58 PFVDGQGPVFRGRIQGANINDQANTGIDGLAGWRVASSQETLNVPVTF----FGKSTLP 113
DB 414 G-TQGTGGGLFGN-----TAQNTGATG--GGLFGNQQTQQQPASTGFGFGTQQQP 463
QY 114 A---GTFTAT 120
DB 464 ATGGGLFGNT 473

RESULT 9

A39199
endoglucanase B (EC 3.2.1.1) - Cellulomonas fimi
C:Species: Cellulomonas fimi
C:Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 09-Jul-2004
C:Accession: A39199
R:Meinke, A.; Braun, C.; Gilkes, N.R.; Kilburn, D.G.; Miller Jr., R.C.; Warren, R.A.J
J. Bacteriol. 173, 308-314, 1991
A:Title: Unusual sequence organization in CenB, an inverting endoglucanase from Cellu
A:Reference number: A39199; MUID:91100298; PMID:1987122
A:Accession: A39199
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1045 <MEI>
A:Cross-references: UNIPROT:P26225; UNIPARC:UPI000012BD04; GB:M64644; NID:g144415; PII
C:Keywords: glycosidase; hydrolase
P:945-1045/Domain: bacterial cellulose-binding domain homology <BCB>
P:946-1044/Disulfide bonds: #status predicted

Query Match 12.5%; Score 82.5; DB 2; Length 1045;
Best Local Similarity 21.1%; Pred. No. 24;
Matches 37; Conservative 21; Mismatches 52; Indels 65; Gaps 5;
QY 1 AAQNTTSANWSQDPGFTGPAVAAGQKVTLSITATG----- 36
DB 545 ASDVTLSANVSECCAQSGKVSAGGTLYGVVELSCVQDIIHFGQSQHRRRQRLTGPAG 604
QY 37 --PHNSVSTAGKASVSGVATVPVVDGQ-----GQPVPRG-RI 72
DB 605 WRPANDPSTGLTQTALAKASAITLYDGSTLVWKSPTGTTTTPPTTGTPTVATGVT 664
QY 73 QGANINDQANT---GIDGLAGWRVASSQETLNVPVTFGKSTLPAGTFTATFV 123
DB 665 VQASLSWAASDAGSGVAGVELRVQGTQTLL-----VGTTTAAAYI 706

RESULT 10

B2954
conserved hypothetical protein PA5544 [imported] - Pseudomonas aeruginosa (strain PAO)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B2954
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: E82954
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-674 <STO>
 A;Cross-references: UNIPROT:Q9HT30; UNIPARC:UPI00000C6066; GB:A8004091; NID:10984043
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: PA5544

Query Match 12.3%; Score 81.5; DB 2; Length 674;
 Best Local Similarity 28.3%; Pred. No. 18;
 Matches 32; Conservative 17; Mismatches 51; Indels 13; Gaps 6;
 QY 15 GFTGPAVAAQKQVTLITATGPHNSVSIAGKAS--VSGGVATVPFVDGQ-QPVFRGR 71
 Db 207 GLFGHRVGGPAKVSFSSALWG-----TITSGVANVVTGQFTPLMKRFGYKPAFAGG 261
 QY 72 IQG-ANINDQANTGIDGLAGRWVASSOETLNVPVTFKSTL-PAGFTATFY 122
 Db 262 VEATASGSQLPPVPMGAVAFIMA---ETINVPYVEIAKALIPALLYFGSVY 311

RESULT 11
 A43291
 collagen alpha 2(I) chain precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: A43291; A54328
 R;Phillips, C.L.; Morgan, A.L.; Lever, L.W.; Wenstrup, R.J.
 Genomics 13, 1345-1346, 1992
 A;Title: Sequence analysis of a full-length cDNA for the murine pro alpha 2(I) collagen
 A;Reference number: A43291; MUID:92372043; PMID:1505972
 A;Accession: A43291
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-1373 <PHI>
 A;Cross-references: UNIPROT:Q01149; UNIPARC:UPI000004DC6; GB:X58251; NID:950488; PIDN:Q01149
 A;Note: sequence extracted from NCBI backbone (NCBI:P:112027)
 R;Phillips, C.L.; Lever, L.W.; Pinnell, S.R.; Quarles, L.D.; Wenstrup, R.J.
 J. Invest. Dermatol. 97, 980-984, 1991
 A;Title: Construction of a full-length murine Proalpha2(I) collagen cDNA by the polymerase chain reaction
 A;Reference number: A54328; MUID:92084969; PMID:1748823
 A;Accession: A54328
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-110 <PH2>
 A;Cross-references: UNIPARC:UPI0000173BA1
 C;Genetics:
 A;Gene: COL1A2
 C;Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology
 C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
 F;1145-1373/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 12.3%; Score 81.5; DB 1; Length 1373;
 Best Local Similarity 34.1%; Pred. No. 40;
 Matches 29; Conservative 9; Mismatches 30; Indels 17; Gaps 4;
 QY 14 PGFTGPAVAAQKQVTLITATGPHNSVSIAGKAS-----VSGGVATVPFVDGQGPVF 68
 Db 957 PGSIGPTGAAG-----APGPHGSGVPAGKGNRGPAGSGVPVGVAVGPRGSGP 1007
 QY 69 RGRIOG--ANINDQANTGIDGLAGW 91
 Db 1008 QG-IRGDKGPGDKGRGLPLGLKY 1031

RESULT 12
 A70524
 probable PPE protein - *Mycobacterium tuberculosis* (strain H37RV)
 C;Species: *Mycobacterium tuberculosis*
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C;Accession: A70524
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: A70524
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-2204 <COL>
 A;Cross-references: UNIPARC:UPI00000C1558; GB:Z96800; GB:AL123456; NID:93261800; PIDN:UPI00000C1558
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: PPE

Query Match 12.3%; Score 81.5; DB 2; Length 2204;
 Best Local Similarity 24.8%; Pred. No. 67;
 Matches 33; Conservative 18; Mismatches 51; Indels 31; Gaps 7;
 QY 17 TGPV---AAGK-----VGTLSITATGPHNSVSIAGKASVS---GGVATVPFVDGQGP 65
 Db 367 TGDIVLSSIAQRAHFGPITIPNITVVGFTTVAIGGPTAITITGGGAIRIPLISIPAA 426
 QY 66 PVFRGRIQGANINDOA---NTGIDGLAG-----WRVASSOETLN--VPVTTFGK 109
 Db 427 PGF---GNSTNPSSGPFNTGAGSGFGNFGGANSFGFWNLASATSGASGLLNVALGS 482
 QY 110 STLPGATFTATFY 122
 Db 483 GLANVTTVSGFY 495

RESULT 13
 S28634
 adhesin AIDA-I precursor - *Escherichia coli* plasmid pIB6
 C;Species: *Escherichia coli*
 C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
 C;Accession: S28634; S22680; S28881; S72657
 R;Benz, I.
 submitted to the EMBL Data Library, March 1992
 A;Reference number: S28634
 A;Accession: S28634
 A;Molecule type: DNA
 A;Residues: 1-1286 <BEN>
 A;Cross-references: UNIPROT:Q03155; UNIPARC:UPI0000125732; EMBL:X65022; NID:942254; PIDN:UPI0000125732
 R;Benz, I.; Schmidt, M.A.
 Mol. Microbiol. 6, 1539-1546, 1992
 A;Title: AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* serotype O157:H7
 A;Reference number: S22680; MUID:92326638; PMID:1625582
 A;Accession: S22680
 A;Molecule type: DNA
 A;Residues: 839-1286 <BE2>
 A;Cross-references: UNIPARC:UPI000017AA2E; EMBL:X65022
 A;Experimental source: strain 2787
 A;Accession: S28881
 A;Molecule type: protein
 A;Residues: 50-56 <BE3>
 A;Cross-references: UNIPARC:UPI000017AA2F
 A;Experimental source: strain 2787
 R;Suhr, M.; Benz, I.; Schmidt, M.A.
 Mol. Microbiol. 22, 31-42, 1996
 A;Title: Processing of the AIDA-I precursor: removal of the outer membrane
 A;Reference number: S72657; MUID:97055419; PMID:8899706
 A;Accession: S72657
 A;Molecule type: protein
 A;Residues: 847-856 <SUH>
 A;Cross-references: UNIPARC:UPI000017AA30
 A;Experimental source: DAEC strain 2787
 C;Genetics:
 A;Genome: plasmid pIB6

C;Keywords: membrane protein

F;1-49/Domain: signal sequence #status predicted <SIG>
F;50-1286/Product: adhesin AIDA-I #status predicted <MAT>

Query Match 12.0%; Score 79.5; DB 2; Length 1286;
Best Local Similarity 24.6%; Pred. No. 55;
Matches 41; Conservative 19; Mismatches 54; Indels 53; Gaps 7;
QY 4 NTSANWQDPQFTGPA-----VAAGQKVGTL-----SITATGPHNS- 40
Db 493 NFLTAVSMFPGTASGANVNLSGRINAFAGNVVGTILNQEGRYVYSGATATSTVGNNEG 552
QY 41 -----VSIAGKGSVSGGVATVPFV-DGQGQVPF-RGRIOGANINDQANTG 84
Db 553 REYVLSGGITDGTVLNSGGLQAVSSGKASATVINEGGAQFPYVDGGQVGTGNIKNGGTI- 611
QY 85 IDGLAGWRVASSQETLNPVTFGK-----STLPAGTFTATFYVQQ 125
Db 612 -----RVDSGASALNIALSSGGLFTTGTATLPELTMAALSVSQ 651

RESULT 14

B99985

Hypothetical protein SA1766 [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C;Accession: B99985

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A;Reference number: B99758; MUID:21311952; PMID:11418146

A;Accession: B99985

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1509 <KUR>

A;Cross-references: UNIPROT:Q99ST6; UNIPARC:UPI00000C7F8E; GB:BA0000018; PID:g13701750; E

A;Experimental source: strain N315

C;Genetics:

A;Gene: SA1766

Query Match 12.0%; Score 79.5; DB 2; Length 1509;
Best Local Similarity 28.6%; Pred. No. 65;
Matches 28; Conservative 9; Mismatches 34; Indels 27; Gaps 3;

QY 9 NWSQDPGFTGPAVAGQKVGTLST--ATGPH-----NSVS 42
Db 1194 NFSKSP-SGTMVKPGDVVGLTGTGTFGLHFEMERNGRHFDPPEYLRNKKGRLS 1252

QY 43 IAGKGSVSGGVATVPFVDGQGPVFRGRIOGANINDQ 80

Db 1253 IGGGATSGSGCATYASRVIRQAQSILGGRYKGIHQ 1290

RESULT 15

G97016

comE-like protein, Metallo beta-lactamase superfamily hydrolase, secreted [imported] - C

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C;Accession: G97016

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: G97016

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-320 <KUR>

A;Cross-references: UNIPROT:Q9YKH3; UNIPARC:UPI00000D46B5; GB:AE001437; PIDN:AAK78922.1;

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC0946

C;Superfamily: Bacillus halodurans late competence protein homolog

Query Match 11.9%; Score 78.5; DB 2; Length 320;
Best Local Similarity 28.0%; Pred. No. 14;
Matches 26; Conservative 13; Mismatches 41; Indels 13; Gaps 2;
QY 17 TGPVAAAGQKVGTLSTATGPHNSVSIAGKGSVSGGVATVPFVDGQGPVFRGRIOGAN 76
Db 142 TNPVVGSTFKVGSANCIVVGPINTNS-----ENLNTYSIVIKLTYGNNKFMFTGDAQSSN 196
QY 77 INDQANTGID-----GLAGWRVASSQETLN 101
Db 197 EPDMINAGFDLSADVLKVGHHGSHSTSSQDFLN 229

Search completed: April 26, 2006, 17:04:59

Job time : 17 secs